



# The Structure of Unstable Power Systems

Joseph M. Abdou

## ► To cite this version:

| Joseph M. Abdou. The Structure of Unstable Power Systems. 2009. halshs-00392515

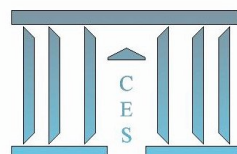
**HAL Id: halshs-00392515**

**<https://shs.hal.science/halshs-00392515>**

Submitted on 8 Jun 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**The Structure of Unstable Power Systems**

Joseph ABDOU

**2009.42**



# The Structure of Unstable Power Systems

Joseph M. Abdou \*

May 13, 2009

**Abstract.** A power system is modeled by an interaction form, the solution of which is called a settlement. By stability we mean the existence of some settlement for any preference profile. Like in other models of power structure, instability is equivalent to the existence of a cycle. Structural properties of the system like maximality, regularity, superadditivity, subadditivity and exactness are defined and used to determine the type of instability that may affect the system. A Stability Index is introduced. Loosely speaking this index measures the difficulty of the emergence of configurations that produce a deadlock. As applications we have a characterization of solvable game forms, an analysis of the structure of their instability and a localization of their stability index in case where solvability fails.

**Keywords:** Interaction Form, Effectivity Function, Stability Index, Nash Equilibrium, Strong Equilibrium, Solvability, Acyclicity, Nakamura Number, Collusion.

**JEL Classification:** C70, D71 **AMS Classification:** 91A44

---

\*Ecole d'Economie de Paris, Paris 1, CNRS, 106-112 boulevard de l'Hôpital 75647 Paris Cedex 13 - France; email: [abdou@univ-paris1.fr](mailto:abdou@univ-paris1.fr)

## Introduction

Stability is an essential requirement for political systems; however it is known that most political institutions are unstable. In this paper we study the structure of instability of power systems. A system enters into instability when it is submitted to contradictory forces that prevent any outcome from being established in such a way that commonly accepted institutions work normally. However, no matter how harmful it may be deemed for the viability of collectivities, instability is not necessarily a course of chaotic actions and events that obey no law: the main thesis of this paper is that there are patterns of instability.

Many countries present the property of being politically split over two main issues. History and geography are accountable for this bipolarity. The main issues can be of socioeconomic type, or of ethnic or religious type. Almost all Western countries are divided between left and right, conservative and liberal, democrats and republicans. Many Middle-Eastern societies are split into pro-Western and anti-Western coalitions. *Bipolarity of opinions does not necessarily translate into instability.* In most Western countries (“democracies”), where governance is based on a written constitution and elections, rules of government are immune to that crude type of instability. Lawmakers strive to define constitutions that avoid instability generated by any bipolar split. By contrast some countries in the Middle-East did experience recently this type of instability. Some Pro-Western coalition formally constituted the ruling power, nevertheless it could not force any significant outcome. The anti-Western coalition itself could oppose any outcome but could force none: Political analysts express this situation by the vocable “political stalemate” or “impasse”.

Some countries, though immune to bipolar instability, could experience more sophisticated types of instability. Many parties with distinct political agendas exist simultaneously. Legal institutions work correctly and choose some ruling coalition with some program. The ruling coalition includes two or more parties who agree on most issues. But the exercise of power becomes impossible when there is a disagreement within the coalition over the implementation of some important issue. Some party in the opposition proposes an alliance to some component of the ruling coalition. As a result, the ruling coalition will eventually be overthrown, and new elections will be held. This scenario may repeat itself. Lawmakers designed institutions that are immune against bipolar stalemates but the political and sociological structure is more complex. Instability may occur. It is important from the point of view of political science to distinguish between this type of instability and the bipolar stalemate.

In this paper, using game theoretic tools, we wish to shed some light on

the structure of instability.

Our approach to instability is similar to that of [5]. In that paper the scope was limited to coalitional power distributions. In the present paper we extend our study so as to encompass the major features of strategic interactions ruled by standard solutions e.g. Nash equilibrium or strong Nash equilibrium. For that purpose we adopt a model of power that is similar to that of effectivity structures introduced by Abdou and Keiding [7]. Here, the general concept is that of *interaction form* and the solution is called a *settlement*.

We consider power systems with abstract outcomes rather than an interaction with specified preferences. This approach allows for the study of systems as such (institutions), where the profile of actual agents is drawn from an arbitrary population. Therefore this article can be viewed as a study in political or social engineering.

In our approach, there is no loss in focusing on the power structure rather than the strategic form. As it turns out, an interaction form can be associated to any strategic form and an equilibrium concept, in such a way that, given any preference profile, an outcome is a settlement of the interaction form if and only if it is an equilibrium outcome of the strategic game form. However other interaction forms exist that are not derived from any strategic game form. An interaction form is said to be *stable* if any preference profile gives rise to some settlement.

Now assume that the interaction form is not stable as it is indeed often the case in political life. As one of the advantages of the model adopted in this paper, a comparison between different power systems is possible within the same framework and therefore, at least in some cases, the model suggests why some systems are deemed more stable than others. This question is most relevant for political institutions, like constitutions or protocols of government formation.

The general idea in studying unstable systems is to obtain some typology of instability. An instability type would determine the general features along which instability is likely to emerge and consequently the lines along which the society is likely to split. Technically, this idea leads to a dissection of the interaction form so as to obtain a graded sequence of substructures. Some properties that describe stability and instability, like maximally, superadditivity and subadditivity are already known, but they are involved in the lowest part of the interaction form, the part that one can identify as the “effectivity function”. Substructures of higher degrees can also be involved in producing instability. For each degree an exactness property is defined: it describes whether the joint action (interaction arrays) of some coalition structure is stronger or not than the independent actions of the coalitions

that compose that structure. The absence of exactness at some degree implies the opportunity of collusion and for maximal interactive forms it signals the existence of some cycle.

The same idea applied to local effectivity functions has led to the definition of an index [5]. Here we extend that definition to interaction forms. The stability index is a number that may be set to infinity in case of stability and that measures the difficulty to block the system, that is to prevent any settlement from prevailing. If this number is low, for instance two, then a simple split in the society with strong opposition power on each side can lead, at polarized preferences, to a stalemate. If the index is high then unless agents possess some intricate preference profile, a settlement can be reached. The index plays a role similar to that of the Nakamura Number for simple games (Nakamura 1979), the difference being that the Nakamura number is defined on the winning coalition structure only, whereas the stability index depends on the whole interaction form. In [5], it is shown that, in case of instability, the index of a maximal local effectivity function is either 2 or 3. Here, even for maximal interaction forms, the index can take any value between 2 and the cardinal of the alternative set. If the interaction form is not  $r$ -exact for some degree  $r$ , then the upper bound on the index is  $r + 2$ .

The paper is organized as follows. In section 1 we define the interaction form as the main object of our model. In section 2 some families of interaction forms are presented. Section 3 is devoted to stability and acyclicity. In section 4 we clarify the notion of cycle and we define the stability index. Section 5 is the most technical part of the paper; it is devoted to the structure of instability. In section 6 we apply our theory to strategic game forms and standard equilibrium notions. In section 7 we provide some examples.

## 1 Interaction forms

Our basic model in this study is that of interaction form. The latter is a power model that can be viewed as a generalization of effectivity functions. An effectivity function is to a game form what a coalitional game is to a strategic game. Effectivity functions play an important role in implementation theory [13, 14, 6, 16, 18, 17, 20, 21]. In other directions, effectivity functions have been used as theoretical tools to analyze solvability problems of strategic game forms. They proved to be particularly relevant for Nash solvability of two-player game forms as well as that of rectangular game forms [9, 10, 11, 2, 3]. However characterizing strong Nash solvability, even for two-player game forms, necessitates the introduction of other effectivity structures [2, 7]. Interaction form can thus be viewed as the general effectivity structure adapted to the solvability problem in the context of any one-shot (pure) equilibrium concept. But the need for such a structure may be independently seen during a careful analysis of any complex interaction

in a political context.

Let  $N$  be a set of agents (also called players, individuals) and let  $A$  be a set of states (also called alternatives, issues). Our approach is founded on the general idea that given some prevailing state, agents dispose of some power to oppose that state, that is to disrupt it if they have an interest to do so. We shall illustrate informally our arguments with an example in politics by taking the case of a government formation in some State. Players are the atomic entities that are endowed with an autonomous wish and will. In our example players may be parties, influential groups, lobbies, etc... A coalition is a subset of players with coordinated action. A government (to be formed) is any element of  $A$ . We start by considering the so-called simple game model. In a simple game, power is withheld by some set  $\mathcal{W}$  of winning coalitions: precisely, a winning coalition has the power to oppose any current or proposed scenario and to propose any other, whereas a losing coalition can oppose no government at all. Whether a winning coalition will object and act in consequence depends on its actual preferences. If no objection is formulated, the government is adopted. In technical terms the outcome is in the core of the simple game given the preferences.

Though simple games can fairly model some decision mechanisms, like weighted majority voting in some institutions, it is too simple to describe political issues underlying a government formation in most countries. This is because in that model a coalition is either absolutely powerful or totally powerless. An effectivity function  $E$  allows a more general distribution of power among the coalitions. In our example, if  $B \in E(S)$  where  $B$  is a subset of possible governments and  $S$  is a coalition, then the latter can upset  $a$  by threatening to form some other government in  $B$  but does not have enough power to force precisely one alternative. However here too, the power of a coalition does not depend on the current state  $a$ , in other words the model takes into account for any coalition, only the part of power that is common in all states. This is a significant restriction in the model since we think that in most interactions, a coalition may achieve something if the current situation is  $a$  and something else if it were  $b$ . One solution is to allow for an effectivity power depending on the state  $a$ . This is the local effectivity function (see [2, 4, 7]). The idea to introduce a power description that is conditional on the state goes back to Rosenthal [19].

Furthermore, it is crucial to note that the same outcome (social state) may be implemented in various ways. For instance, assume that  $(S_1, S_2, S_3)$  are involved in government formation, and that many scenarios may lead to  $a$ , then in one scenario  $S_2$  can upset  $a$  by proposing  $B_2$ , while in the other coalition  $S_3$  can upset the same  $a$  by proposing  $B_3$ . Therefore we may consider the active coalition structure  $(S_2, S_3)$  as the opposition actor. Although the action is not coordinated among the coalitions of the active

coalition structure, there is at least some collusion to oppose  $a$ . Taking into account this idea compels us to introduce the concept of interaction form. In an interaction form we take into account (1) the dependence of the interaction power on the actual state in all possible scenarios that lead to that state and (2) the joint opposition power that the outcome may activate. In order to block (or dismantle) some outcome, the joint action of many coalitions may be necessary. Such a joint power potentially activated at some state is called an interaction array. An interaction form describes all possible interaction arrays available at some state. Technically, interaction forms are to game forms and equilibria what effectivity functions are to game forms and the core. But interaction forms can be studied abstractly without any reference to game forms. This was first done in [2] and later on in [4, 7].

Given a preference profile for players, a state can potentially awaken any interaction array available at that state. Whether there is some incentive to upset (or dismantle) that state depends on the preference profile. A state is called a settlement at that profile if no such incentive exists. The interaction form is said to be stable if it admits at least one settlement at any preference profile.

An interaction form can be associated to any strategic game form together with a given equilibrium concept. The model that we adopt in this paper is similar to that of effectivity structures introduced by Abdou and Keiding [7]. The advantages of the present model are (1) that it allows for the representation of various equilibrium concepts within the *same* interaction form, whereas the other one is specific to one equilibrium concept, (2) that in the current model, operations like projections *faithfully* reflect the change in the underlying active coalition structure. The unifying aspect of this model lies in the fact that any one-shot solution (e.g., Nash equilibrium, strong Nash equilibrium,  $\beta$ -core) applied to some game form coincides with the settlement set of the associated interactive form. One of the advantages of working in the general framework of power models than that of strategic models, is that solvability for any standard solution can be expressed by the same nice though somehow difficult condition namely *acyclicity*. Any failure in stability can be seen as the effect of some generalized Condorcet cycle. An interaction form can thus be viewed as an intrinsic representation of the power inherent to the system ruled by some equilibrium concept. Once the interaction form is obtained, only the settlement set matters. This unification, a byproduct of our model, is in the same spirit as the one obtained in Greenberg's *theory of social situations* [8] even though our purpose remains different. On the one hand, a general theory of interactions is not within the scope of our paper; on the other hand, the model that we adopt applies to game forms rather than games because our aim is to study properties of power systems (institutions) and not properties of specific interactions.



## 1.1 Notations

Throughout this paper we shall consider a finite set  $N$  the elements of which are called players or agents and a finite set  $A$  the elements of which are called alternatives or states. We make use of the following notational conventions: For any set  $D$ , we denote by  $\mathcal{P}(D)$  the set of all subsets of  $D$  and by  $\mathcal{P}_0(D) = \mathcal{P}(D) \setminus \{\emptyset\}$  the set of all non-empty subsets of  $D$ . Elements of  $\mathcal{P}_0(N)$  are called *coalitions*. If  $S \in \mathcal{P}_0(N)$  then  $N \setminus S$  is denoted  $S^c$ . Similarly if  $B \in \mathcal{P}(A)$ ,  $A \setminus B$  is denoted  $B^c$ .  $L(A)$  will denote the set of all linear orders on  $A$  (that is all binary relations on  $A$  which are complete, transitive, and antisymmetric). If  $R \in L(A)$ , and  $a, b \in A, a \neq b$ ,  $a R b$  means that  $a$  is preferred to  $b$  in the linear order  $R$ . We also denote by  $R^\circ$  the strict relation associated to  $R$ . A preference profile (over  $A$ ) is a map from  $N$  to  $L(A)$ , so that a preference profile is an element of  $L(A)^N$ . For every preference profile  $R_N \in L(A)^N$  and  $S \in \mathcal{P}_0(N)$  we put

$$P(a, S, R_N) = \{b \in A \mid b R_i^\circ a, \forall i \in S\}$$

(so that  $P(a, S, R_N)$  consists of all the outcomes considered to be better than  $a$  by all members of the coalition  $S$ ), and  $P^c(a, S, R_N) = A \setminus P(a, S, R_N)$ .

## 1.2 The model

In order to model a power distribution, we define an object called *interaction form* in the same vein as the *effectivity structure* of Abdou and Keiding [7]. The elements of  $A$  are viewed as (social, political) situations or states. At any state  $a \in A$  we dispose of a description of the acting power of the agents in the society. This acting power which depends generally on  $a$  is represented by a set of interaction arrays. If the state of the society is  $a$ , some individuals or coalitions can move or threaten to move to other states upsetting therefore the state  $a$ . Thus power is described as a multipolar force that can be used to upset a status quo. Formally we define the following:

**Definition 1.1** An *interaction array* on  $(N, A)$  is a mapping  $\varphi : \mathcal{P}_0(N) \rightarrow \mathcal{P}(A)$ . An interaction array  $\varphi$  is said to be *simple* if there exists  $(B^1, \dots, B^n) \in \mathcal{P}(A)^n$  such that for all  $S \in \mathcal{P}_0(N)$ :  $\varphi(S) = \cap_{i \in S} B^i$ . Such an interaction array is said to be associated to  $(B^1, \dots, B^n)$ .

Let  $\Phi \equiv \Phi(N, A)$  be the set of all interaction arrays. We endow  $\Phi$  with the partial order  $\leq$  where  $\varphi \leq \varphi'$  if and only if  $\varphi(S) \subset \varphi'(S)$  for all  $S \in \mathcal{P}_0(N)$ . By *active coalition structure* (ACS hereafter)  $\mathcal{M}$  we mean any subset of  $\mathcal{P}_0(N)$ . By federation  $\mathfrak{F}$ , we mean any subset of ACS. The *support* of  $\varphi$  denoted  $[\varphi]$  is the active coalition structure formed by all coalitions  $S \in \mathcal{P}_0(N)$  such that  $\varphi(S) \neq \emptyset$ . We denote by  $\Phi_0 \equiv \Phi_0(N, A)$  the subset of interaction arrays with non empty support. The range of  $\varphi$  is the set

$\rho(\varphi) := \cup_{S \in \mathcal{P}_0(N)} \varphi(S)$ . More generally for any ACS  $\mathcal{M}$ , the range of  $\varphi$  in  $\mathcal{M}$  is the set  $\rho_{\mathcal{M}}(\varphi) := \cup_{S \in \mathcal{M}} \varphi(S)$ .

**Definition 1.2** An *interaction form* over  $(N, A)$  is a mapping  $\mathcal{E}$  from  $\mathcal{P}_0(A)$  to subsets of  $\Phi_0$  satisfying the following conditions:

- (i)  $\varphi \leq \varphi', \varphi \in \mathcal{E}[U] \Rightarrow \varphi' \in \mathcal{E}[U]$
- (ii)  $U \subset V \Rightarrow \mathcal{E}[V] \subset \mathcal{E}[U]$

An interaction form  $\mathcal{E}$  is said to satisfy the *sheaf property* if for all  $U, V \in \mathcal{P}_0(A)$ :

$$\mathcal{E}(U \cup V) = \mathcal{E}(U) \cap \mathcal{E}(V).$$

$\mathcal{E}$  is said to be *locally effective* if for all  $U \in \mathcal{P}_0(A)$ ,  $\varphi \in \Phi_0$ :

$$U \subset \rho(\varphi) \Rightarrow \varphi \in \mathcal{E}[U]$$

$\mathcal{E}$  is said to be *standard* if it has the sheaf property and is locally effective.

We may think of an interaction array in  $\mathcal{E}[U]$  as a description of an available move of the agents given any state in  $U$ . To interpret the statement  $\varphi \in \mathcal{E}[\{a\}]$ , one has to assume that  $a$  may occur in different scenarios that are not explicit in the model; any scenario leading to state  $a$  may arouse some coalition  $S$  that have the power to drive the outcome into  $\varphi(S)$ . In an interaction array all such potential “moves” are described. If, for instance,  $\varphi(S)$  and  $\varphi(T)$  are the only nonempty components of  $\varphi$ , then for some situations with outcome  $a$ , coalition  $S$  has the power to reach  $\varphi(S)$  and for some situations with the same outcome, coalition  $T$  has the power to reach  $\varphi(T)$ . Within each coalition action is coordinated (as in any coalitional game), but there is no coordination between  $S$  and  $T$ . Our model is universal in the sense that *a priori* any coalition may react to some state in  $U$ . Nevertheless, the fact that  $\varphi(S) = \emptyset$  for some  $S$  means that coalition  $S$  is inhibited or deactivated and therefore that the power represented by  $\varphi$  holds without the participation of  $S$ . Therefore the support of  $\varphi$  is in fact the *active* coalition structure behind  $\varphi$ .

Whether coalitions have a real incentive to make their move depends on the actual preferences. This is why we introduce the following:

**Definition 1.3** Let  $R_N \in L(A)^N$ . An alternative  $a$  is *dominated* at  $R_N$  if there exists some  $U \in \mathcal{P}_0(A)$ ,  $U \ni a$ , and some  $\varphi \in \mathcal{E}(U)$  such that  $\varphi(S) \subset P(a, S, R_N)$  for all  $S \in \mathcal{P}_0(N)$ . The alternative  $a$  is a *settlement* at  $R_N$  if it is not dominated at  $R_N$ . The set of all settlements at  $R_N$  will be denoted:  $\text{Stl}(\mathcal{E}, R_N)$ .

It follows from the definition that, even if there is no coordination of actions between coalitions, some *collusion* may exist between them (see subsection 5.5). For instance, if for  $\varphi \in \mathcal{E}[\{a\}]$ ,  $\varphi(S)$  and  $\varphi(T)$  are nonempty and if,

given the preference profile, both coalitions wish to oppose  $a$ , they can do it: actually it will be  $S$  that oppose the outcome in some scenarios and  $T$  in some others. This is why  $a$  cannot survive. The absence of a settlement at some preference profile can be expressed as an impasse or a deadlock. Stability, therefore, is a highly desirable property for an interaction form.

We end this subsection by a useful definition. For any  $\varphi : \mathcal{P}_0(N) \rightarrow \mathcal{P}(A)$  we put:  $\varphi^*(S) = \varphi(S)^c$ , and for an mapping  $\mathcal{E}$  from  $\mathcal{P}_0(A)$  to subsets of  $\Phi$  we define:

$$\mathcal{E}^*[U] = \{\varphi \in \Phi \mid \varphi^* \notin \mathcal{E}[U]\} \quad (1)$$

An *interaction coform* over  $(N, A)$  is a mapping  $\mathcal{E}$  from  $\mathcal{P}_0(A)$  to subsets of  $\Phi$  such that  $\mathcal{E}^*$  is an interaction form. All our theory could have been formulated using interaction coforms instead of interaction forms. The two approaches are mathematically equivalent. Interaction forms highlight the opposition power (that is the  $\beta$ -version of power). They describe the power to upset some state while coforms describe the conservation power that forces that state (that is the  $\alpha$ -version of power). Those are the two faces of the same power system, but since we focus on destabilizing power we prefer the  $\beta$ -version.

## 2 Some families of interaction forms

In this section we present some generic ways to define interaction forms. Any game form ruled by a one-shot solution concept gives rise to an interaction form, the settlement of which reflects faithfully that solution. Similarly an effectivity function with the core as solution gives rise to some interaction form to which it can be identified. It follows that the model of interaction forms imbeds the main strategic aspects of game forms as well as the cooperative aspects of coalitional forms.

### 2.1 Interaction form and strategic game form

In this subsection, starting from a strategic game form and an equilibrium concept, we derive a description of the underlying power distribution, thus defining some interaction form. This derivation of the power embedded in a strategic form follows the same pattern as the derivation of the  $\alpha$ - and the  $\beta$ -effectivity functions [13] and more recently the derivation of the more general effectivity structures [2, 7].

Let  $G = (X_1, \dots, X_n, A, g)$  be a strategic game form. The set of players is  $N = \{1, \dots, n\}$ ,  $X_i$  is the strategy set of players  $i$ ,  $g : \prod_{i \in N} X_i \rightarrow A$  is the outcome function, assumed to be surjective. For each preference profile  $R_N \in L(A)^N$ , the game form  $G$  induces a game  $(X_1, \dots, X_n; Q_1, \dots, Q_n)$

with the same strategy spaces as in  $G$  and with the  $Q_i$  is the preorder on  $X_N$  defined by:  $x_N Q_i y_N$  if and only if  $g(x_N) R_i g(y_N)$  for  $x_N, y_N \in X_N$ . We denote this game by  $G(R_N)$ .

Let  $\mathcal{M}$  be an active coalition structure. A strategy array  $x_N \in X_N$  is an  $\mathcal{M}$ -equilibrium of the game  $G(R_N)$  if there is no coalition  $S \in \mathcal{M}$  and  $y_S \in X_S$  such that  $g(y_S, x_{S^c}) R_i^\circ g(x_N)$  for all  $i \in S$ .

An alternative  $a$  is an  $\mathcal{M}$ -equilibrium outcome of  $G$  at  $R_N$  if there exists some equilibrium  $x_N \in X_N$  of  $G(R_N)$  such that  $g(x_N) = a$ . Denote by  $EO_{\mathcal{M}}(G, R_N)$  the set of all equilibrium outcomes of  $(G, R_N)$ .

The game form  $G$  is said to be *solvable in  $\mathcal{M}$ -equilibrium* or  $\mathcal{M}$ -solvable, if for each preference profile  $R_N \in L(A)^N$ , the game  $G(R_N)$  has an  $\mathcal{M}$ -equilibrium. In particular, when  $\mathcal{M} = \mathcal{N} \equiv \{\{1\}, \dots, \{n\}\}$ , an  $\mathcal{M}$ -equilibrium is a Nash equilibrium. Similarly, when  $\mathcal{M} = \mathcal{P}_0(N)$ , an  $\mathcal{M}$ -equilibrium is a strong Nash equilibrium.

The *interaction form* associated with  $G$  and  $\mathcal{M}$  is the mapping  $\mathcal{E}_\beta^G(\mathcal{M})$  defined as follows: For  $U \in \mathcal{P}_0(A)$ :

$$\mathcal{E}_\beta^G(\mathcal{M})[U] = \{\varphi \in \Phi_0(N, A) \mid \forall y_N \in g^{-1}(U), \exists S \in \mathcal{M}, \exists x_S \in X_S : g(x_S, y_{S^c}) \in \varphi(S)\} \quad (2)$$

**Lemma 2.1** *Let  $G$  be a game form. For any active coalition structure  $\mathcal{M}$ , the set of  $\mathcal{M}$ -equilibrium outcomes of  $G$  at  $R_N$  coincides with the settlement set of  $\mathcal{E}_\beta^G(\mathcal{M})$  at  $R_N$ . Therefore  $G$  is  $\mathcal{M}$ -solvable if and only if  $\mathcal{E}_\beta^G(\mathcal{M})$  is stable.*

There are other ways to associate interaction forms to some game form, namely by using other solutions that cannot be formulated directly as an  $\mathcal{M}$ -equilibrium. This is the case for the local core and the  $\beta$ -core.

An alternative  $a$  is in the *local core* of  $G$  at  $R_N$  if there is no coalition  $S \in \mathcal{P}_0(N)$  with the following property : for any  $z_N \in X_N$  such that  $g(z_N) = a$  there exists  $y_S \in X_S$  such that  $g(y_S, z_{S^c}) R_i^\circ g(x_N)$  for all  $i \in S$ . Denote by  $C_1(G, R_N)$  the local core of  $G$  at  $R_N$ .

An alternative  $a$  is in the  $\beta$ -core of  $G$  at  $R_N$  if there is no coalition  $S \in \mathcal{P}_0(N)$  with the following property: for any  $z_N \in X_N$ , there exists  $y_S \in X_S$  such that  $g(y_S, z_{S^c}) R_i^\circ g(x_N)$  for all  $i \in S$ . Denote by  $C_\beta(G, R_N)$  or  $C_0(G, R_N)$ , the  $\beta$ -core of  $G$  at  $R_N$ .

We may introduce interaction forms  $\mathcal{E}_1^G$  and  $\mathcal{E}_0^G$  associated to the local core and the  $\beta$ -core respectively as follows: For any  $U \in \mathcal{P}_0(A)$ , we define the following:

$$\mathcal{E}_1^G[U] = \{\varphi \in \Phi \mid \exists S \in \mathcal{P}_0(N), \forall y_N \in g^{-1}(U), \exists x_S \in X_S : g(x_S, y_{S^c}) \in \varphi(S)\} \quad (3)$$

$$\mathcal{E}_0^G[U] = \mathcal{E}_1^G[A] \quad (4)$$

Remark that, in general, there is no  $\mathcal{M} \subset \mathcal{P}_0(N)$  such that  $\mathcal{E}_1^G = \mathcal{E}_\beta^G(\mathcal{M})$ . Nevertheless we have:

$$\mathcal{E}_1^G[U] = \cup_{S \in \mathcal{P}_0(N)} \mathcal{E}_\beta^G(\{S\})[U].$$

(See subsection 5.1 for more details about projections and restrictions).

**Lemma 2.2** *The local core (resp.  $\beta$ -core) of  $G$  at  $R_N$  coincides with the settlement set of  $\mathcal{E}_1^G$  (resp.  $\mathcal{E}_0^G$ ) at  $R_N$ . Therefore  $G$  is locally stable (resp. stable) if and only if  $\mathcal{E}_1^G$  (resp.  $\mathcal{E}_0^G$ ) is stable.*

All the interaction forms defined in this way are standard (definition 1.2).

Corresponding to any solution concept, one can introduce interaction coforms. For instance corresponding to  $\mathcal{M}$ -equilibrium one can define the *interaction coform* associated with  $G$  and  $\mathcal{M}$  as the mapping  $\mathcal{E}_\alpha^G(\mathcal{M})$  where for all  $U \in \mathcal{P}_0(A)$ :

$$\mathcal{E}_\alpha^G(\mathcal{M})[U] = \{\psi \in \Phi_0(N, A) \mid \exists y_N \in g^{-1}(U), \forall S \in \mathcal{M}, \forall x_S \in X_S : g(x_S, y_{S^c}) \in \psi(S)\} \quad (5)$$

For any  $x = x_N \in X_N$  define the interaction array  $\psi_x$  by  $\psi_x(S) = g(x_{S^c}, X_S)$  ( $S \in \mathcal{P}_0(N)$ ). In applications it is often easier to compute the family  $(\psi_x, x \in X_N)$ .  $\psi_x$  is said to be a basic interaction array at  $g(x)$ . Clearly  $\mathcal{E}_\alpha^G(\mathcal{M}) = \mathcal{E}_\beta^G(\mathcal{M})^*$  and  $\mathcal{E}_\beta^G(\mathcal{M}) = \mathcal{E}_\alpha^G(\mathcal{M})^*$ . It is also possible, starting from the basic interaction arrays, to compute  $\mathcal{E}_\alpha^G(\mathcal{M})$  and  $\mathcal{E}_\beta^G(\mathcal{M})$ :

$$\mathcal{E}_\alpha^G(\mathcal{M})[U] = \{\varphi \in \Phi \mid \exists x \in g^{-1}(U), \forall S \in \mathcal{M} : \psi_x(S) \subset \varphi(S)\} \quad (6)$$

$$\mathcal{E}_\beta^G(\mathcal{M})[U] = \{\varphi \in \Phi \mid \forall x \in g^{-1}(U), \exists S \in \mathcal{M} : \psi_x(S) \cap \varphi(S) \neq \emptyset\} \quad (7)$$

In what follows we illustrate these notions with two examples.

**Example 2.3** (Majority game form with status quo) Let  $N = \{1, 2, 3\}$  and  $A = \{a, b, c\}$ . We consider the majority game form  $G$  with status quo. Formally  $X_1 = X_2 = X_3 = A$  and  $g(x_1, x_2, x_3) = x$  whenever  $|\{i \in N \mid x_i = x\}| \geq 2$  or  $g(x_1, x_2, x_3) = a$  and  $\{x_1, x_2, x_3\} = A$ . Let  $\mathcal{M} = \mathcal{P}_0(N)$ . We can compute  $\mathcal{E}_\alpha^G \equiv \mathcal{E}_\alpha^G(\mathcal{M})$  by listing for each  $z \in A$  all  $\psi_x$  where  $x \in X_1 \times X_2 \times X_3$  is such that  $g(x) = z$ .

The list for  $z = b$  is given by the table 2.3 where rows 2 to 7 correspond to arrays  $\psi_x$  ( $x \in g^{-1}(b)$ ) (precisely row 2 represents  $\psi_{b,b,b}$ , row 3 represents  $\psi_{c,b,b}$ , etc ...). The columns list the coalitions (1 for  $\{1\}$ , 12 for  $\{1, 2\}$ , etc

...).

An interaction array  $\psi$  is in  $\mathcal{E}_\alpha^G[b]$  if and only if it contains some  $\psi_x$  listed in the table. An interaction array  $\varphi$  is in  $\mathcal{E}_\beta^G(\mathcal{M})[b]$  if and only if for each line  $\psi_x$ , there corresponds a column  $S$  such that  $\varphi(S)$  intersects  $\psi_x(S)$ . At the last line of the table we represent one instance of such a  $\varphi$  (at each line  $x$  we affected an upper bar, in some entry  $(x, S)$ , on the alternative that is in  $\psi_x(S)$  and  $\varphi(S)$ ).

$x_1$	$x_2$	$x_3$	1	2	3	12	13	23	123
$b$	$b$	$b$	$b$	$b$	$b$	$abc$	$abc$	$abc$	$abc$
$c$	$b$	$b$	$b$	$abc$	$abc$	$abc$	$abc$	$abc$	$abc$
$b$	$c$	$b$	$abc$	$b$	$abc$	$abc$	$abc$	$abc$	$abc$
$b$	$b$	$c$	$abc$	$abc$	$b$	$abc$	$abc$	$abc$	$abc$
$a$	$b$	$b$	$b$	$\bar{a}b$	$\bar{a}b$	$abc$	$abc$	$abc$	$abc$
$b$	$a$	$b$	$ab$	$b$	$\bar{a}b$	$abc$	$abc$	$abc$	$abc$
$b$	$b$	$a$	$\bar{a}b$	$ab$	$b$	$abc$	$abc$	$abc$	$abc$
$\varphi$			$a$	$a$	$a$	$\emptyset$	$b$	$c$	$\emptyset$

Figure 1: Table of the basic interaction arrays of  $\mathcal{E}_\alpha^G[b]$

$\mathcal{E}_\alpha^G[c]$  is obtained from  $\mathcal{E}_\alpha^G[b]$  by a permutation of the roles of  $b$  and  $c$ . Table 2.3 does the same job for  $z = a$ .

$x_1$	$x_2$	$x_3$	1	2	3	12	13	23	123
$a$	$a$	$a$	$a$	$a$	$a$	$abc$	$abc$	$abc$	$abc$
$a$	$a$	$b$	$ab$	$ab$	$a$	$abc$	$abc$	$abc$	$abc$
$a$	$b$	$a$	$ab$	$a$	$ab$	$abc$	$abc$	$abc$	$abc$
$b$	$a$	$a$	$a$	$ab$	$ab$	$abc$	$abc$	$abc$	$abc$
$a$	$a$	$c$	$ac$	$ac$	$a$	$abc$	$abc$	$abc$	$abc$
$a$	$c$	$a$	$ac$	$a$	$ac$	$abc$	$abc$	$abc$	$abc$
$c$	$a$	$a$	$a$	$ac$	$ac$	$abc$	$abc$	$abc$	$abc$
$a$	$b$	$c$	$abc$	$ac$	$ab$	$abc$	$abc$	$abc$	$abc$
$a$	$c$	$b$	$abc$	$ab$	$ac$	$abc$	$abc$	$abc$	$abc$
$b$	$a$	$c$	$ac$	$abc$	$ab$	$abc$	$abc$	$abc$	$abc$
$c$	$a$	$b$	$ab$	$abc$	$ac$	$abc$	$abc$	$abc$	$abc$
$b$	$c$	$a$	$ac$	$ab$	$abc$	$abc$	$abc$	$abc$	$abc$
$c$	$b$	$a$	$ab$	$ac$	$abc$	$abc$	$abc$	$abc$	$abc$

Figure 2: Table of the basic interaction arrays of  $\mathcal{E}_\alpha^G[a]$

We shall see (subsection 5.1) that from  $\mathcal{E}_\beta^G$  one can extract  $\mathcal{E}_\beta^G(\mathcal{M})$  for each active coalition structure.

**Example 2.4** (Unanimity game form). Let  $X$  be a finite set such that  $|X| \geq 2$  and  $0 \notin X$  and let  $A = X \cup \{0\}$ . The  $n$ -player *unanimity game form* on  $X$  is defined by setting:  $X_i = X, i = 1, \dots, n$  and

$$\begin{aligned} g(x_1, \dots, x_n) &= x \quad \text{if } x_1 = \dots = x_n = x \\ g(x_1, \dots, x_n) &= 0 \quad \text{otherwise.} \end{aligned}$$

Let  $\mathcal{M} = \mathcal{P}_0(N)$  and  $a \in X$ . One has:

$$\begin{aligned} \mathcal{E}_\alpha^G(\mathcal{M})[a] &= \{\psi \in \Phi \mid \psi(N) = A \text{ and } \forall S \neq N : \{0, a\} \subset \psi(S)\} \\ \mathcal{E}_\beta^G(\mathcal{M})[a] &= \{\varphi \in \Phi \mid \varphi(N) \neq \emptyset \text{ or } \exists S \neq N : \{0, a\} \cap \varphi(S) \neq \emptyset\} \end{aligned}$$

The formula for  $\mathcal{E}_\alpha^G(\mathcal{M})[0]$  is more complex. Later on (see example 6.10) we shall compute  $\mathcal{E}_\alpha^G(\mathcal{M})[0]$  for some particular ACS  $\mathcal{M}$ .

## 2.2 Effectivity functions as interaction forms

**Definition 2.5** An *effectivity function* on  $(N, A)$  is a mapping  $E : \mathcal{P}(N) \rightarrow \mathcal{P}(\mathcal{P}_0(A))$  such that :

- (i)  $E(\emptyset) = \emptyset$ ,
- (ii)  $B \in E(S), B \subset B' \Rightarrow B' \in E(S)$

The core of  $E$  at  $R_N$  denoted  $C(E, R_N)$  is the set of outcomes  $a \in A$  such that there is no coalition  $S \in \mathcal{P}_0(N)$  such that  $P(a, S, R_N) \in E(S)$ . We put:

$$\mathbf{E} = \{\varphi \in \Phi \mid \exists S \in \mathcal{P}_0(N) : \varphi(S) \in E(S)\} \quad (8)$$

The “canonical ” interaction form associated to  $E$  is defined by:  $\mathcal{E}[U] = \mathbf{E}$  for all  $U \in \mathcal{P}_0(A)$ . For future use we mention the following properties that play some role in the study of stability. An effectivity function  $E$  is said to be:

*monotonic w.r.t. players* if for all  $S, T \in \mathcal{P}_0(N)$ ,

$$S \subset T \Rightarrow E(S) \subset E(T), \quad (9)$$

*regular* if for all  $S_1 \in \mathcal{P}_0(N), S_2 \in \mathcal{P}_0(N)$ ,

$$S_1 \cap S_2 = \emptyset, B_1 \in E(S_1), B_2 \in E(S_2) \Rightarrow B_1 \cap B_2 \neq \emptyset, \quad (10)$$

*maximal* if for all  $S \in \mathcal{P}_0(N), B \in \mathcal{P}_0(A)$ ,

$$B^c \notin E(S^c) \Rightarrow B \in E(S), \quad (11)$$

*superadditive* if for all  $S_1 \in \mathcal{P}_0(N), S_2 \in \mathcal{P}_0(N)$ ,

$$S_1 \cap S_2 = \emptyset, B_1 \in E(S_1), B_2 \in E(S_2) \Rightarrow B_1 \cap B_2 \in E(S_1 \cup S_2), \quad (12)$$

*subadditive* if for all  $S_1 \in \mathcal{P}_0(N), S_2 \in \mathcal{P}_0(N)$ ,

$$B_1 \cap B_2 = \emptyset, B_1 \in E(S_1), B_2 \in E(S_2) \Rightarrow B_1 \cup B_2 \in E(S_1 \cap S_2). \quad (13)$$

### 3 Stability and acyclicity

Let  $\mathcal{E}$  be an interaction form. Cycles appear naturally in the study of interaction forms stability (see [12, 6] for cycles in effectivity functions and [7] for cycles in more general effectivity structures). For any  $\varphi \in \Phi$  and  $i \in N$  we put:  $\rho^i(\varphi) = \cup_{S \ni i} \varphi(S)$  (the range of player  $i$  in  $\varphi$ ).

**Definition 3.1** An  $\mathcal{E}$  - family is any  $r$ -tuple  $((U_1, \varphi_1), \dots, (U_r, \varphi_r))$  where:  $U_k \in \mathcal{P}_0(A)$ ,  $\varphi_k \in \mathcal{E}[U_k]$  ( $k = 1, \dots, r$ ). An  $\mathcal{E}$  -family is a *cycle* in  $\mathcal{E}$  if it satisfies :

- (i)  $\cup_{k=1}^r U_k = A$ ,
- (ii) For any  $i \in N$  and  $\emptyset \neq J \subset \{1, \dots, r\}$  there exists  $k \in J$  such that for all  $l \in J$  :  $U_k \cap \rho^i(\varphi_l) = \emptyset$ .

The covering  $(U_1, \dots, U_r)$  will be called the *basis* of the cycle. The natural number  $r$  is the *order* (or the *length*) of the cycle. Such a cycle will be called an  $r$ - cycle. A cycle is said to be *simple* if for all  $k = 1, \dots, r$ ,  $\varphi_k$  is simple (definition 1.1) The interaction form  $\mathcal{E}$  is said to be *acyclic* if it has no cycles.

**Remark 3.2** *Relation to the literature.* In [7] the authors define a model similar to the present one and called effectivity structure. Let  $\mathcal{E}$  be an interactive form. Then one can associate to  $\mathcal{E}$  an effectivity structure say  $\mathcal{E}^-$  as follows: for any  $U \in \mathcal{P}_0(A)$ , the  $n$ -tuple  $(B^1, \dots, B^n) \in \mathcal{P}(A)^n$  is in  $\mathcal{E}^-[U]$  if and only if the simple interaction array associated to  $(B^1, \dots, B^n)$  is in  $\mathcal{E}[U]$ . The notion of cycle for an effectivity structure is defined in such a way that to any cycle in  $\mathcal{E}^-$  the associated sequence of simple interaction arrays is a cycle in  $\mathcal{E}$  and conversely. Moreover, by proposition 4.3 (section 4)  $\mathcal{E}$  has a cycle of order  $r$  if and only if  $\mathcal{E}^-$  has a cycle of the same order. The following Theorem, that provides a necessary and sufficient condition for stability of interaction game forms, can thus be deduced from theorem 6 of [7].

**Theorem 3.3**  $\mathcal{E}$  is stable if and only if  $\mathcal{E}$  is acyclic.

### 4 Cycles and stability index

The existence of a cycle is equivalent to instability. As it has been argued in the introduction, instability that may occur in a political system is not a chaotic or unanalyzable matter. In a country governed by perennial institutions, unstable situations often present the same features. Governments are toppled almost in the same way. Institutions are generally paralyzed along the same opposition lines. That clearly indicates some flaw in the current institutions or the governance system. In this study this flaw has been identified as the existence of some cycle. Studying the structure of instability



amounts to characterizing the nature of cycles. However one must admit that cycles are complex objects. Therefore we start by a clarification of the notion of cycle.

#### 4.1 Cycles: the combinatorial formulations

**Proposition 4.1** *In definition 3.1, one can replace condition (ii) by any of the two following conditions:*

(iib) *For any  $i \in N$  and  $\emptyset \neq J \subset \{1, \dots, r\}$  there exists  $l \in J$  such that for all  $k \in J : U_k \cap \rho^i(\varphi_l) = \emptyset$ ,*

(iic) *For any  $i \in N$  there exists a permutation  $k_1, \dots, k_r$  of  $\{1, \dots, r\}$  such that  $[(U_{k_1} \cup \dots \cup U_{k_j})] \cap [\rho^i(\varphi_{k_j}) \cup \dots \cup \rho^i(\varphi_{k_r})] = \emptyset$ .*

*Proof.* Assume that (ii) is satisfied. We shall prove (iic). Let  $i \in N$ . By taking  $J = \{1, \dots, r\}$  in (ii), we can choose  $k_1 \in \{1, \dots, r\}$  such that  $U_{k_1} \cap (\cup_{l \in \{1, \dots, r\}} \rho^i(\varphi_l)) = \emptyset$ . Assume that  $k_1, \dots, k_j$  have been chosen such that:  $[(U_{k_1} \cup \dots \cup U_{k_j})] \cap [\cup_{l \in \{1, \dots, r\} \setminus \{k_1, \dots, k_{j-1}\}} \rho^i(\varphi_l)] = \emptyset$ . By taking  $J = \{1, \dots, r\} \setminus \{k_1, \dots, k_j\}$  we can choose  $k_{j+1}$ . Thus (iic) is proved by induction.

Now we prove (iic)  $\Rightarrow$  (ii) and (iib). Assume that (iic) is satisfied. Let  $\emptyset \neq J \subset \{1, \dots, r\}$ . If we take  $k = k_s$  where  $s = \min\{j \mid k_j \in J\}$ , then (ii) is satisfied by that choice of  $k$ . If we take  $l = k_t$  where  $t = \max\{j \mid k_j \in J\}$ , then (iib) is satisfied by that choice of  $l$ .  $\square$

Let  $r \in \mathbb{N}$ ,  $r \geq 1$ , let  $\Sigma_r$  be the set of all selections of  $\mathcal{P}_0(\{1, \dots, r\})$ . Precisely  $\theta \in \Sigma_r$  if  $\theta : \mathcal{P}_0(\{1, \dots, r\}) \rightarrow \{1, \dots, r\}$  and  $\theta(J) \in J$  for all  $J \in \mathcal{P}_0(\{1, \dots, r\})$ . To any  $(\varphi_1, \dots, \varphi_r) \in \Phi^r$ ,  $\theta \equiv (\theta^1, \dots, \theta^n) \in (\Sigma_r)^n$  and  $k \in \{1, \dots, r\}$  we associate:

$$\mathcal{J}_k^{i, \theta} := \{J \in \mathcal{P}_0(\{1, \dots, r\}) \mid \theta^i(J) = k\} \quad (14)$$

$$\mathcal{R}_k^{i, \theta}(U_1, \dots, U_r) := A \setminus \bigcup_{J \in \mathcal{J}_k^{i, \theta}} \bigcup_{l \in J} U_l \quad (15)$$

$$\mathcal{A}_k^\theta(\varphi_1, \dots, \varphi_r) := \bigcup_{i=1}^n \bigcup_{J \in \mathcal{J}_k^{i, \theta}} \bigcup_{l \in J} \rho^i(\varphi_l) \quad (16)$$

In what follows  $\mathcal{A}_k^\theta(\varphi_1, \dots, \varphi_r)$  will be denoted  $\mathcal{A}_k^\theta$  and  $\mathcal{R}_k^{i, \theta}(U_1, \dots, U_r)$  will be denoted  $\mathcal{R}_k^{i, \theta}$ .

**Proposition 4.2** *An  $\mathcal{E}$ -family  $((U_1, \varphi_1), \dots, (U_r, \varphi_r))$  is a cycle if and only if  $\cup_{k=1}^r U_k = A$  and if there exists  $\theta = (\theta^1, \dots, \theta^n) \in (\Sigma_r)^n$  such that for all  $k \in \{1, \dots, r\}$  one has :  $U_k \cap \mathcal{A}_k^\theta = \emptyset$ .*

*Proof.* Let  $((U_1, \varphi_1), \dots, (U_r, \varphi_r))$  be a cycle with basis  $(U_1, \dots, U_r)$ . For each  $i \in N$  we define  $\theta^i \in \Sigma_r$  as follows: By property (ii) of definition 3.1, we put  $\theta^i(J) = k$  if for all  $l \in J$ , we have  $U_k \cap \rho^i(\varphi_l) = \emptyset$ . It follows that

$U_k \cap \mathcal{A}_k^\theta = \emptyset$  ( $k = 1 \dots, r$ ). Conversely if these relations hold then property (ii) of definition 3.1 is clearly satisfied.  $\square$

In the following we clarify the relation between cycles and simple cycles (definition 3.1).

**Proposition 4.3** *Let  $((U_1, \varphi_1), \dots, (U_r, \varphi_r))$  be an  $\mathcal{E}$ -family such that  $(U_1, \dots, U_r)$  be a covering of  $A$ . The following are equivalent:*

- (i)  $((U_1, \varphi_1), \dots, (U_r, \varphi_r))$  is a cycle,
- (ii) There exists  $\theta = (\theta^1, \dots, \theta^n) \in (\Sigma_r)^n$  such that for all  $k \in \{1, \dots, r\}$  the simple interaction array  $\psi_k$  associated to  $(\mathcal{R}_k^{1, \theta}, \dots, \mathcal{R}_k^{n, \theta})$  satisfies  $\varphi_k \leq \psi_k$  ( $k = 1 \dots, r$ ),
- (iii) There exist simple interaction arrays  $\psi_1, \dots, \psi_r$  such that  $\varphi_k \leq \psi_k$  ( $k = 1 \dots, r$ ) and  $((U_1, \psi_1), \dots, (U_r, \psi_r))$  is a cycle.

*Proof.* (iii)  $\Rightarrow$  (i) is straightforward. We first prove (i)  $\Rightarrow$  (ii). Let  $((U_1, \varphi_1), \dots, (U_r, \varphi_r))$  be a cycle and let  $i \in N$ . In view of (iib) of proposition 4.1, one can find  $\theta^i \in \Sigma_r$  such that for all  $J \in \mathcal{P}_0(\{1, \dots, r\})$ ,  $\rho^i(\varphi_{\theta^i(J)}) \cap (\cup_{l \in J} U_l) = \emptyset$ . It follows that for any  $k \in \{1, \dots, r\}$  one has  $\rho^i(\varphi_k) \cap (\cup_{J \in \mathcal{I}_k^i} \cup_{l \in J} U_l) = \emptyset$  or equivalently  $\rho^i(\varphi_k) \subset \mathcal{R}_k^{i, \theta}$ . Define a simple interaction array  $\psi_k$  by  $\psi_k(S) = \cap_{i \in S} \mathcal{R}_k^{i, \theta}$  ( $S \in \mathcal{P}_0(N)$ ). Clearly for all  $S \in \mathcal{P}_0(N)$  one has  $\varphi_k(S) \subset \psi_k(S)$  or equivalently  $\varphi_k \leq \psi_k$ .

Now we prove (ii)  $\Rightarrow$  (iii). If the  $\psi_k$  are as in (ii),  $\psi_k \in \mathcal{E}[U_k]$  and  $\rho^i(\psi_k) = \mathcal{R}_k^{i, \theta}$  so that property (iib) of proposition 4.1 is satisfied for  $((U_1, \psi_k), \dots, (U_r, \psi_r))$  entailing that the latter is indeed a cycle.  $\square$

The following shows that existence of a cycle of some order is concomitant to the existence of a simple cycle of the same order. This result, combined with remark 3.2 proves in particular, that an interaction form is acyclic if and only if its associated effectivity structure is acyclic.

**Corollary 4.4** *A covering  $(U_1, \dots, U_r)$  of  $A$  is the basis of some cycle of  $\mathcal{E}$  if and only if there exists  $\theta = (\theta^1, \dots, \theta^n) \in (\Sigma_r)^n$  such that for all  $k \in \{1, \dots, r\}$  the simple interaction array  $\psi_k$  associated to  $(\mathcal{R}_k^{1, \theta}, \dots, \mathcal{R}_k^{n, \theta})$  is in  $\mathcal{E}(U_k)$ .*

Finally one can formulate the problem of the existence of a cycle as a fixed point theorem. Let  $\mathcal{C}_r$  be the set of coverings  $(U_1, \dots, U_r)$  of  $A$ . Let  $\Psi^{(r)} \equiv (\Psi_1^{(r)}, \dots, \Psi_r^{(r)}) : (\Sigma_r)^n \times \mathcal{C}_r \rightarrow \Phi^r$  be the map where for all  $k$   $\Psi_k^{(r)}(\theta^1, \dots, \theta^n; U_1, \dots, U_r)$  denotes the simple interaction array associated to  $(\mathcal{R}_k^{1, \theta}, \dots, \mathcal{R}_k^{n, \theta})$ . Then there exists some cycle of  $\mathcal{E}$  of order  $r$  if and only if there exists  $\theta \in (\Sigma_r)^n$  and  $(U_1, \dots, U_r) \in \mathcal{C}_r$  such that:

$$\Psi_k^{(r)}(\theta^1, \dots, \theta^n; U_1, \dots, U_r) \in \mathcal{E}[U_k] \quad (k = 1, \dots, r) \quad (17)$$

The ingredients that compose a cycle are thus: a covering (or a partition)  $(U_1, \dots, U_r)$  and a combinatorial setting  $\theta = (\theta^1, \dots, \theta^n) \in (\Sigma_r)^n$ . A typology of cycles is an equivalence relation on the set of cycles: two cycles are in the same class if and only if they are “homologous” in some sense. In this paper we limit the study to the simplest equivalence relation, namely the one where two cycles are considered as equivalent if and only if they are of the same order. This number called the index, measures, in a given interaction form, the complexity attached to the emergence of a cycle. This is only a first step in the understanding of instability. We implicitly assume that cycles of higher order are more difficult to emerge. The stability index is an indicator for the likelihood of the potential occurrence of instability.

**Definition 4.5** The *stability index* of  $\mathcal{E}$ , denoted  $\sigma(\mathcal{E})$ , is the minimal order of a cycle in  $\mathcal{E}$ . This number is set to  $+\infty$  if  $\mathcal{E}$  is acyclic.

## 4.2 Index and merging alternatives

Let  $f : A \rightarrow A'$  be a map. If  $\varphi' \in \Phi(N, A')$  we denote  $f^{-1} \circ \varphi'$ , the element  $\varphi$  of  $\Phi(N, A)$  defined by  $\varphi(S) = (f^{-1} \circ \varphi')(S)$  for all  $S \in \mathcal{P}_0(N)$ . For any interaction form  $\mathcal{E}$  on  $(N, A)$  we define the interaction form  $\mathcal{E}^f$  on  $(N, A')$  as follows: For  $U' \in \mathcal{P}_0(A')$  :

$$\mathcal{E}^f[U'] = \{\varphi' \in \Phi(N, A') \mid f^{-1} \circ \varphi' \in \mathcal{E}[f^{-1}(U')]\}$$

The  $\mathcal{E}^f$ - family  $((U'_1, \varphi'_1), \dots, (U'_r, \varphi'_r))$ , is a cycle of  $\mathcal{E}^f$  if and only if the  $\mathcal{E}$ - family  $((f^{-1}(U'_1), f^{-1} \circ \varphi'_1), \dots, (f^{-1}(U'_r), f^{-1} \circ \varphi'_r))$  is a cycle. It follows that if  $\mathcal{E}$  is acyclic then  $\mathcal{E}^f$  is acyclic.

Let  $((U_1, \varphi_1), \dots, (U_r, \varphi_r))$  be a cycle of  $\mathcal{E}$  based on the partition  $(U_1, \dots, U_r)$ . Let  $A'$  be some set with  $r$  elements  $A' := \{u_1, \dots, u_r\}$  and let  $f : A \rightarrow A'$  be defined by  $f(a) = u_k$  for  $a \in U_k$ . Define  $\varphi' \in \Phi(N, A')$  by putting  $\varphi'_k(S) := f(\varphi_k(S))$  ( $S \in \mathcal{P}_0(N)$ ). For any  $S \in \mathcal{P}_0(N)$  and  $k, l \in \{1, \dots, r\}$  one has  $U_k \cap \varphi_l(S) = \emptyset$  if and only if  $\{u_k\} \cap f(\varphi_l(S)) = \emptyset$ . It follows that  $((\{u_1\}, \varphi'_1), \dots, (\{u_r\}, \varphi'_r))$  is a cycle of  $\mathcal{E}^f$  based on the partition  $(\{u_1\}, \dots, \{u_r\})$ . Therefore we have the following characterization of the index, that generalizes a similar result obtained for local effectivity functions ([5], Theorem 4.3).

**Theorem 4.6** *The index of an unstable interaction form  $\mathcal{E}$  is the smallest integer  $s$  for which there exists a surjective mapping  $f : A \rightarrow \{1, \dots, s\}$  such that  $\mathcal{E}^f$  is unstable.*

This characterization provides an interpretation of the stability index. Assume that an interaction form is unstable with a stability index  $\sigma$ , then merging some social states (or alternatives) results in a decrease of the number of alternatives and a transformation of the interaction form in a way that

respects power distribution. This is the interpretation of the transformation  $\mathcal{E} \rightarrow \mathcal{E}^f$ . This transformation may occur, for instance, when the agents do not distinguish any more between two previously different alternatives. If the number of the new alternatives is less than  $\sigma$  then the new interaction form will be stable. When  $\sigma = 2$  (see special subsection 5.6 on bipolarity) alternatives can be partitioned into two aggregates, or two major issues, over which the society can be opposed, and the power of agents or institutions allowed by the rules is such that either issue can be opposed and neither can be forced. Efficient institutions are designed in order to avoid this situation. Most modern political systems, though immune to bipolarity, may suffer higher order cycles. On the one hand, the formation of a cycle of lower order is less costly than that of a cycle of higher order. This is due to the cooperation factor (coalitions must coordinate their strategies) and the collusion factor (many coalitions may be needed to contribute in order to upset the same outcome, because multiple scenarios with the same outcome are potentially available). On the other hand the more sophisticated is a society, the more unlikely is the formation of a cycle of lower order: this is because the number of relevant (politically significant) alternatives is high. Compare two societies that have been given the same institutions, with index 3 for instance. The first society is split by an ethnic or religious strife, while the second society lives with many social issues considered as relevant. In the first society, ethnic or religious conflict cannot create an impasse; while it is possible, in the second case, to observe instability, though the occurrence of the latter is tempered by the fact that conditions of cyclicity are somehow difficult to crystallize.

## 5 The structure of instability

In this section we shall try to determine or at least provide an estimation for the stability index. For that purpose we are going to devise some structural properties on interaction forms. Precisely we shall extract some appropriate substructures from the basic structure. The simplest substructure is the one that contains the local and the global effectivity functions  $E_1$  and  $E_0$ , respectively. This will be the object of subsection 5.3. Higher order substructures may also be extracted (subsection 5.4) and the notion of exactness is defined (subsection 5.5) for each. The tools that we provide are not merely technical. Each of the notions and operations that we shall define (projection, federation, exactness) have a theoretical role and possess an intuitive interpretation. We start by the projection operation.

### 5.1 Projection

As explained above, an interaction array is defined on  $\mathcal{P}_0(N)$ , so that the model allows *a priori* the surge of any coalition. Now it may be the case that

institutionally (by law or structural impossibilities) some coalitions are not allowed to form: only coalitions in some  $\mathcal{M} \subset \mathcal{P}_0(N)$  can actually be active. For instance in a legislature (Senate, House of representatives) only some coalitions are practically possible. The definitions can be adapted in order to take into account these institutional limitations. An alternative  $a$  is  $\mathcal{M}$ -dominated at the preference profile  $R_N$  if there exists some  $U \in \mathcal{P}_0(A)$ ,  $U \ni a$  and  $\varphi \in \mathcal{E}(U)$  such that  $[\varphi] \subset \mathcal{M}$  and  $\varphi(S) \subset P(a, S, R_N)$  for all  $S \in \mathcal{M}$ . The alternative  $a$  is an  $\mathcal{M}$ -settlement at  $R_N$  if it is not  $\mathcal{M}$ -dominated at  $R_N$ . The set of all  $\mathcal{M}$ -settlements at  $R_N$  will be denoted:  $\text{Stl}(\mathcal{M})(\mathcal{E}, R_N)$ . In order to cope with this situation without changing our model, we would like to define an interaction form that reflects the activity of the active coalition structure  $\mathcal{M}$ . This is precisely the role of the following operation. For  $\varphi \in \Phi$  the projection of  $\varphi$  on  $\mathcal{M}$ , denoted  $\varphi_{\mathcal{M}}$  is defined by:

$$\varphi_{\mathcal{M}}(S) := \begin{cases} \varphi(S) & \text{if } S \in \mathcal{M} \\ \emptyset & \text{if } S \notin \mathcal{M} \end{cases}$$

The *projection* of  $\mathcal{E}$  on  $\mathcal{M}$  is the interaction form  $\mathcal{E}(\mathcal{M})$  defined by :

$$\mathcal{E}(\mathcal{M})[U] := \{\varphi \in \Phi \mid \varphi_{\mathcal{M}} \in \mathcal{E}[U]\} \quad (18)$$

In particular  $\mathcal{E}(\emptyset)[U] = \emptyset$  and if  $\mathcal{M}_1$  and  $\mathcal{M}_2$  are subsets of  $\mathcal{P}_0(N)$  then  $\mathcal{E}(\mathcal{M}_1)(\mathcal{M}_2) = \mathcal{E}(\mathcal{M}_1 \cap \mathcal{M}_2)[U] = \mathcal{E}(\mathcal{M}_1)[U] \cap \mathcal{E}(\mathcal{M}_2)[U]$ .

Clearly one has:  $\text{Stl}(\mathcal{M})(\mathcal{E}, R_N) = \text{Stl}(\mathcal{E}(\mathcal{M}), R_N)$ .

As a general remark one can see that allowing that some  $\varphi(S)$  be empty is not only an innocuous convention: it has to be interpreted as meaning: the intervention of  $S$  is not needed in the interaction array  $\varphi$ .

## 5.2 Federation

Now consider the case of a legislative body composed of two Chambers. If  $a$ , a confidence motion for instance, is disrupted by an active coalition structure, say  $\mathcal{M}$ , in Chamber 1, then  $a$  is discarded. If  $a$  passes Chamber 1 unopposed then it has to be presented to Chamber 2, where the active structure is  $\mathcal{T}$ . We assume that  $\mathcal{M}$  and  $\mathcal{T}$  exert their sovereignty independently. Then in order to analyze the whole governance structure we need to introduce the federation  $\{\mathcal{M}, \mathcal{T}\}$ .

We shall call *federation* any set  $\mathfrak{F}$  of active coalition structures. Federations will be denoted by symbols  $\mathfrak{F}, \mathfrak{M}, \mathfrak{P} \dots$

Examples of federations are the Congress in the USA (Senate and House of representatives, where some proposal is submitted successively to both legislatures), the Parliament in France (Assemblée Nationale and Sénat). More generally any institution with two or more levels of independent legislatures. In a federation any component is an active coalition structure. It

is assumed that each component acts independently of other components. One can note a difference between the activity of an ACS and that of a federation. For an ACS, in order to rule out some state  $a$ , a joint action of the components-coalitions may be needed: in general no coalition of its own is assumed to have the power to rule out that state. On the federation level, if the state is ruled out by some active structure (e.g., Chamber 1), then it is discarded.

The notions of  $\mathfrak{F}$ -settlement and  $\mathfrak{F}$ -stability are defined in consequence: An alternative  $a$  is  $\mathfrak{F}$ -dominated at the preference profile  $R_N$  if there exists some  $\mathcal{M} \in \mathfrak{F}$ ,  $U \in \mathcal{P}_0(A)$ ,  $U \ni a$  and  $\varphi \in \mathcal{E}(\mathcal{M})[U]$  and  $\varphi(S) \subset P(a, S, R_N)$  for all  $S \in \mathcal{M}$ . The alternative  $a$  is an  $\mathfrak{F}$ -settlement at  $R_N$  if it is not  $\mathfrak{F}$ -dominated at  $R_N$ . The set of  $\mathfrak{F}$ -settlements at  $R_N$  will be denoted:  $\text{Stl}_{\mathfrak{F}}(\mathcal{E}, R_N)$ . The second operation that we need is the following:

The *restriction* of  $\mathcal{E}$  to the federation  $\mathfrak{F}$  is defined by:

$$\mathcal{E}_{\mathfrak{F}}[U] := \bigcup_{\mathcal{M} \in \mathfrak{F}} \mathcal{E}(\mathcal{M})[U] \quad (19)$$

It is clear from the definition that: if  $\mathfrak{F}_1$  and  $\mathfrak{F}_2$  are two federations, then  $\mathcal{E}_{\mathfrak{F}_1 \cup \mathfrak{F}_2}[U] = \mathcal{E}_{\mathfrak{F}_1}[U] \cup \mathcal{E}_{\mathfrak{F}_2}[U]$ .

Moreover one has:  $\text{Stl}_{\mathfrak{F}}(\mathcal{E}, R_N) = \text{Stl}(\mathcal{E}_{\mathfrak{F}}, R_N) = \bigcap_{\mathcal{M} \in \mathfrak{F}} \text{Stl}(\mathcal{E}(\mathcal{M}), R_N)$ .

**Example 5.1** a) Let  $\mathcal{S} = \mathcal{P}_0(N)$ .  $\mathcal{S}$  is the active coalition structure corresponding to the situation where all coalitions have some joint power. A settlement for  $\mathcal{S}$  is similar to a strong Nash equilibrium outcome ( see subsection 2.1).

b) Let  $\mathcal{M} = \mathcal{N} := \{\{i\} | i \in N\}$ .  $\mathcal{N}$  is the active coalition structure corresponding to the situation where only individuals have some joint power. A settlement for  $\mathcal{N}$  is similar to a Nash equilibrium outcome (subsection 2.1). In the case of interaction forms associated to strategic game forms  $\mathcal{E}_{\alpha}^G(\mathcal{N})$  is the projection of  $\mathcal{E}_{\alpha}^G(\mathcal{P}_0(N))$  on  $\mathcal{N}$ .

c) Let  $\mathfrak{F} \equiv \mathfrak{M}_1 = \{\{S\} | S \in \mathcal{P}_0(N)\}$ .  $\mathfrak{M}_1$  is the federation where every active coalition structure is a single coalition. This is a context where every coalition have an independent (as opposed to joint) power. A settlement in this case is similar to an element of the local core (subsection 2.1). In the case of interaction forms associated to strategic game forms, with solution the local core, the corresponding interaction form  $\mathcal{E}_1^G$  is the restriction of  $\mathcal{E}_{\alpha}^G(\mathcal{P}_0(N))$  on  $\mathfrak{M}_1$ .

**Remark 5.2** 1) Let  $\mathfrak{M} = \{\{S\} : S \in \mathcal{M}\}$ . There is an important difference between  $\mathcal{E}(\mathcal{M})$  and  $\mathcal{E}_{\mathfrak{M}}$ .  $\mathcal{E}(\mathcal{M})$  reflects the uncoordinated power of one active coalition structure acting jointly and simultaneously, whereas  $\mathfrak{M}$  reflects the aggregate power of coalitions acting independently of each other.

2) What is the result of adding a coalition  $S$  to some active coalition structure  $\mathcal{M}$  ? We obtain the active coalition structure  $\mathcal{M}' = \mathcal{M} \cup \{S\}$ . What is the result if we federate  $\{S\}$  to  $\mathcal{M}$  ? We obtain the federation  $\mathfrak{F} = \{\mathcal{M}, \{S\}\}$ . In general, since  $\mathcal{M} \subset \mathcal{M}'$  and  $\{S\} \subset \mathcal{M}'$  we have  $\mathcal{E}_{\mathfrak{F}}[U] = \mathcal{E}(\mathcal{M})[U] \cup \mathcal{E}(\{S\})[U] \subset \mathcal{E}(\mathcal{M}')[U]$ . Now what happens if  $S = N$  ? If we assume that any  $\varphi$  with  $\varphi(N) \neq \emptyset$  is in  $\mathcal{E}[U]$  ( $N$  is sovereign), then we have  $\mathcal{E}(\mathcal{M}') = \mathcal{E}_{\mathfrak{F}}$ .

One of the advantages of our present model of interaction forms compared to the one proposed in Abdou and Keiding [7] is that it allows for projections and restrictions in a way that will prove relevant in the case of interaction forms derived from strategic game forms. In subsection 2.1, we associated to any game form  $G$  and ACS  $\mathcal{M}$  an interaction form denoted  $\mathcal{E}_{\beta}^G(\mathcal{M})$ . Then it is clear from the definition that  $\mathcal{E}_{\beta}^G(\mathcal{M})$  is the projection of  $\mathcal{E}_{\beta}^G(\mathcal{P}_0(N))$  on  $\mathcal{M}$  or using the notations of the present subsection  $\mathcal{E}_{\beta}^G(\mathcal{M}) = (\mathcal{E}_{\beta}^G(\mathcal{P}_0(N)))(\mathcal{M})$ . By restriction of  $\mathcal{E}_{\beta}^G$  to the federation  $\mathfrak{M}_1$  we have the local effectivity function of  $G$  namely  $\mathcal{E}_1^G$  etc... In some sense  $\mathcal{E}_{\beta}^G(\mathcal{P}_0(N))$  is “universal”. The model of effectivity structure of [7] does not allow for projections and restrictions. The following subsections make precise this idea and extend it to any interactive form.

### 5.3 Induced effectivity: $\mathcal{E}_0, \mathcal{E}_1, E_1, E_0$

In subsection 2.2 we have shown how interaction forms generalize effectivity functions; now we show how they induce effectivity functions.

If  $S \in \mathcal{P}(N)$  and  $B \in \mathcal{P}_0(A)$ , denote by  $\delta_{S,B}$  the element of  $\Phi$  such that:

$$\delta_{S,B}(T) = \begin{cases} \emptyset & \text{if } T \neq S \\ B & \text{if } T = S \end{cases}$$

Let  $\mathcal{E}$  be an interaction form and let  $\mathfrak{M}_1 := \{\{S\} \mid S \in \mathcal{P}_0(N)\}$  be the federation composed of all ACS that are singletons.

Let  $\mathcal{E}_1$  be the restriction of  $\mathcal{E}$  to  $\mathfrak{M}_1$ , let  $\mathbf{E}_0 := \mathcal{E}_1[A]$  and let  $\mathcal{E}_0$  be the interaction form defined by  $\mathcal{E}_0[U] := \mathbf{E}_0$  for all  $U \in \mathcal{P}_0(A)$ .

For  $S \in \mathcal{P}_0(N)$  and  $U \in \mathcal{P}_0(A)$ , we define  $E_1$  and  $E_0$  as follows:

$$E_1[U](S) = \{B \in \mathcal{P}_0(A) \mid \delta_{S,B} \in \mathcal{E}[U]\} \quad (20)$$

$$\begin{aligned} E_0(S) &= \{B \in \mathcal{P}_0(A) \mid \delta_{S,B} \in \mathcal{E}[A]\} \\ &= E_1[A](S) \end{aligned} \quad (21)$$

$E_1$  is the *local effectivity function* induced by  $\mathcal{E}$ ,  $E_0$  is the (*global*) *effectivity function* induced by  $\mathcal{E}$ . Moreover for any  $U \in \mathcal{P}_0(A)$ , one has:

$$\mathcal{E}_1[U] = \{\varphi \in \Phi \mid \exists S \in \mathcal{P}_0(N) : \varphi(S) \in E_1[U](S)\} \quad (22)$$

$$\mathbf{E}_0 = \{\varphi \in \Phi \mid \exists S \in \mathcal{P}_0(N) : \varphi(S) \in E_0(S)\} \quad (23)$$

Let  $C(E_0, R_N)$  be the core of  $E_0$  at  $R_N$  ( subsection 2.2). One has  $\text{Stl}(\mathcal{E}_0, R_N) = C(E_0, R_N)$ .

**Definition 5.3** An interaction form  $\mathcal{E}$  is said to be *maximal* (resp. *regular*, *superadditive*, *subadditive*), if  $E_0$  satisfies that property (subsection 2.2).

From now on, we shall use  $E_0$  or  $\mathcal{E}_0$  indifferently when we want to refer to the effectivity function extracted from  $\mathcal{E}$ . We shall also use  $C(E_0, \cdot)$  or  $\text{Stl}(\mathcal{E}_0, \cdot)$  to refer to the core correspondence of  $E_0$ .

## 5.4 Higher order derived structures : $\mathcal{E}_r, r \geq 2$

$\mathcal{E}_0$  (or equivalently  $E_0$ ) will play a fundamental role in the sequel. However  $\mathcal{E}_0$  represents the lowest level in the interaction: the level of *independent* and global power of single coalitions. In order to study the structure of instability especially the stability index, one may need higher level “effectivity”. Therefore we are led to the definition of a graded family of derived structures. Let  $\mathfrak{A}_1$  be the federation  $\mathfrak{M}_1 \equiv \{\{S\} \mid S \in \mathcal{P}_0(N)\}$ . For  $r \geq 2$ , let  $\mathfrak{A}_r$  be the set of active coalition structures  $\mathcal{M}$  such that  $|\mathcal{M}| = r$ ,  $N \notin \mathcal{M}$  and for all  $S, T \in \mathcal{M}$  such that  $S \neq T$  we have  $S \cup T = N$ .

When  $r \geq 2$ ,  $\mathcal{M} \in \mathfrak{A}_r$  if and only if  $\mathcal{M}^* := \{T \mid T^c \in \mathcal{M}\}$  is composed of  $r$  nonempty disjoint subsets of  $N$ . We put  $\mathfrak{M}_r = \cup_{k=1}^{k=r} \mathfrak{A}_k$ . An element of  $\mathfrak{M} \equiv \mathfrak{M}_n$  will be called an *admissible* ACS.

Let  $\mathcal{E}$  be an interaction form. Let  $\mathcal{E}_r \equiv \mathcal{E}_{\mathfrak{M}_r}$  be the restriction of  $\mathcal{E}$  to  $\mathfrak{M}_r$ . One has:

$$\text{Stl}(\mathcal{E}_0, R_N) \supset \text{Stl}(\mathcal{E}_1, R_N) \supset \cdots \supset \text{Stl}(\mathcal{E}_n, R_N) \supset \text{Stl}(\mathcal{E}, R_N) \quad (24)$$

## 5.5 Exactness and opportunity of collusion

In order to obtain some indications on the existence of cycles in some derived substructure  $\mathcal{E}_r$ , we define the notion of exactness as a generalization of the notion bearing the same name introduced in Abdou [2]. In what follows, if  $\varphi \in \Phi$ ,  $\mathcal{M}$  an ACS, we denote by  $\rho_{\mathcal{M}}(\varphi) := \cup_{S \in \mathcal{M}} \varphi(S)$ . We recall that  $E_0$  is the (global) effectivity function derived from  $\mathcal{E}$  (see equation 21). For any active coalition structure  $\mathcal{M}$ , we define the following sets:

$$\Phi_*(\mathcal{M}) = \{\varphi \in \Phi \mid \rho_{\mathcal{M}}(\varphi) \neq A\} \quad (25)$$

$$\mathbf{E}_0(\mathcal{M}) = \{\varphi \in \Phi_*(\mathcal{M}) \mid \exists S \in \mathcal{M}, \varphi(S) \in E_0(S)\} \quad (26)$$

$$\mathbf{E}_*(\mathcal{M}) = \{\varphi \in \Phi_*(\mathcal{M}) \mid \forall a \notin \rho_{\mathcal{M}}(\varphi), \varphi \in \mathcal{E}(\mathcal{M})[A \setminus \rho_{\mathcal{M}}(\varphi)]\} \quad (27)$$

$$\mathbf{E}_{\xi}(\mathcal{M}) = \{\varphi \in \Phi_*(\mathcal{M}) \mid \exists a \notin \rho_{\mathcal{M}}(\varphi), \varphi \in \mathcal{E}(\mathcal{M})[a]\} \quad (28)$$

$$\mathbf{D}(\mathcal{M}) = \{\varphi \in \Phi \mid S, T \in \mathcal{M}, S \neq T \Rightarrow \varphi(S) \cap \varphi(T) = \emptyset\} \quad (29)$$



Remark that we have the equalities:

$$\begin{aligned}\mathbf{E}_0(\mathcal{M}) &= \cup_{a \in A} \{\varphi \mid a \notin \rho_{\mathcal{M}}(\varphi), \varphi \in \mathcal{E}_0(\mathcal{M})[A]\} \\ \mathbf{E}_*(\mathcal{M}) &= \cup_{a \in A} \{\varphi \mid a \notin \rho_{\mathcal{M}}(\varphi), \varphi \in \mathcal{E}(\mathcal{M})[A \setminus \rho_{\mathcal{M}}(\varphi)]\} \\ \mathbf{E}_{\xi}(\mathcal{M}) &= \cup_{a \in A} \{\varphi \mid a \notin \rho_{\mathcal{M}}(\varphi), \varphi \in \mathcal{E}(\mathcal{M})[a]\}\end{aligned}$$

so that:  $\mathbf{E}_0(\mathcal{M}) \subset \mathbf{E}_*(\mathcal{M}) \subset \mathbf{E}_{\xi}(\mathcal{M})$ .

In order to have an intuitive interpretation of these sets, assume for a while that  $\mathcal{E}$  is standard (definition 1.2).  $\mathbf{E}_{\xi}(\mathcal{M})$  represents the power to reach interaction arrays in  $\Phi_*(\mathcal{M})$  that an ACS holds when it acts according to  $\mathcal{E}$ .  $\mathbf{E}_*(\mathcal{M})$  which, for a standard  $\mathcal{E}$ , is equal to  $\mathcal{E}(\mathcal{M})[A] \cap \Phi_*(\mathcal{M})$  represents the global power held by  $\mathcal{M}$ .  $\mathbf{E}_0(\mathcal{M})$  represents the sum of the global power that coalitions in  $\mathcal{M}$  hold independently and separately.

**Definition 5.4** Let  $\mathcal{M}$  be an ACS.  $\mathcal{E}$  is  $\mathcal{M}$ -exact if one has:

$$\mathbf{E}_{\xi}(\mathcal{M}) \cap \mathbf{D}(\mathcal{M}) = \mathbf{E}_0(\mathcal{M}) \cap \mathbf{D}(\mathcal{M}) \quad (30)$$

Let  $r \in \mathbb{N}$ ,  $1 \leq r \leq n$ .  $\mathcal{E}$  is  $r$ -exact if it is  $\mathcal{M}$ -exact for all  $\mathcal{M} \in \mathfrak{M}_r$ . In particular  $\mathcal{E}$  is said to be *exact* if it is 1-exact and *fully exact* if it is  $n$ -exact.

What is the *additional* power that an ACS  $\mathcal{M}$  derives from the interaction compared to the sum of the global power that coalitions hold separately ? If  $\varphi \in \mathbf{E}_{\xi}(\mathcal{M})$  and  $\varphi \notin \mathbf{E}_0(\mathcal{M})$ . Then for some  $a \notin \rho_{\mathcal{M}}(\varphi)$ , we have  $\varphi \in \mathcal{E}(\mathcal{M})[a]$  and  $\varphi \notin \mathcal{E}_0(\mathcal{M})[A]$ . The interactive power of the coalitions of  $\mathcal{M}$  is stronger than their independent power: we say that there is an *opportunity of collusion* at  $a$ . Indeed, if  $a$  is proposed and coalitions of  $\mathcal{M}$  prefer to reject  $a$ , then it may be the case that  $\varphi$  cannot be in their power because no coalition can do it independently since  $\forall S \in \mathcal{M}, \varphi(S) \notin E_0(S)$ , whereas they can very well do it jointly since  $\varphi \in \mathcal{E}(\mathcal{M}[a])$ : that is, in any scenario where  $a$  is proposed some coalition  $S \in \mathcal{M}$  can counter it by  $\varphi(S)$ . Collusion is not cooperation. Collusion (in politics, diplomacy or war) expresses precisely a situation where two or more distinct forces, though not formally cooperating, have an objective interest to target the same state (in our interpretation their common goal is to upset that state).  $\mathcal{M}$ -exactness amounts to say that restricted to  $\mathbf{D}(\mathcal{M})$  there is no such additional power. Therefore when restricted to  $\mathbf{D}(\mathcal{M})$ , the joint power of the ACS  $\mathcal{M}$  can be decoupled, that is to say, distributed between the coalitions that compose the ACS. In the following example 1-exactness fails though the interaction form does not present any local dependence.

**Example 5.5** Let  $|A| = p \geq 3$ . For any  $U \in \mathcal{P}_0(A)$ , put:  $\mathcal{E}[U] =$

$$\{\varphi \in \Phi \mid \varphi(N) \neq \emptyset \text{ or } \exists S, T \neq N, S \cup T = N, |\varphi(S) \cup \varphi(T)| \geq p - 1\}$$

It is clear that  $\mathcal{E}$  does not depend on  $U$ . Moreover for any  $S \in \mathcal{P}_0(N)$   $a \in A$ ,  $S \neq N$ :  $E_1[a](S) = E_0(S) = \{B \mid |B| \geq p-1\}$ ,  $E_0(N) = E_1[a](N) = \mathcal{P}_0(A)$ , and for any  $S, T \neq \emptyset, N$  with  $S \cup T = N$ :

$$\begin{aligned} \mathbf{E}_0(\{S, T\}) &= \{\varphi \mid |\varphi(S) \cup \varphi(T)| \leq p-1\} \cap \{\varphi \mid |\varphi(S)| \geq p-1 \text{ or } |\varphi(T)| \geq p-1\} \\ &= \{\varphi \mid |\varphi(S)| = p-1 \text{ and } \varphi(T) = \emptyset\} \cup \{\varphi \mid |\varphi(T)| = p-1 \text{ and } \varphi(S) = \emptyset\} \\ \mathbf{E}_\xi(\{S, S^c\}) &= \{\varphi \mid |\varphi(S) \cup \varphi(S^c)| = p-1\} \end{aligned}$$

Since there is no dependence on  $U$ ,  $\mathcal{E}$  is 1-exact. However  $\mathcal{E}$  is not 2-exact: indeed  $\mathbf{E}_0(\{S, S^c\}) \cap \mathbf{D}(\{S, S^c\}) \neq \mathbf{E}_\xi(\{S, S^c\}) \cap \mathbf{D}(\{S, S^c\})$ . There is some additional power in  $\mathbf{E}_\xi(\{S, S^c\})$  that exceeds the union of the separate effectivity power of  $S$  and  $S^c$  as represented by  $E_0(S)$  and  $E_0(S^c)$  respectively.

**Example 5.6** Let  $|A| = p \geq 2$ . For any  $a \in A$ ,  $U \in \mathcal{P}_0(A)$ , let  $w_a : \mathcal{P}_0(N) \rightarrow \{1, \dots, p\}$ ,  $\mathcal{E}[a] = \{\varphi \mid \exists S \in \mathcal{P}_0(N) : |\varphi(S)| \geq w_a(S)\}$  and  $\mathcal{E}(U) = \cap_{a \in U} \mathcal{E}[a]$ .

Let  $w(S) = \max_{a \in A} w_a(S)$ , then one obtains:

$$\begin{aligned} E_1[a](S) &= \{B \mid |B| \geq w_a(S)\}, \quad E_0(S) = \{B \mid |B| \geq w(S)\}, \\ \mathbf{E}_\xi(\{S\}) &= \{\varphi \mid \varphi(S) \neq A \text{ and } |\varphi(S)| \geq \min_{a \notin \varphi(S)} w_a(S)\} \end{aligned}$$

Then it is easy to see that  $\mathcal{E}$  is 1-exact if and only if  $\forall a \in A, \forall S \in \mathcal{P}_0(N) : w_a(S) = w(S)$  ( $\mathcal{E}$  is independent of  $U$ ). Assume that this condition is satisfied. For any ACS  $\mathcal{M}$  and  $\varphi \in \Phi$ , with  $\rho_{\mathcal{M}}(\varphi) \neq A$ , one has  $\varphi \in \mathbf{E}_\xi(\mathcal{M})$  if and only if  $\exists S \in \mathcal{M} : |\varphi(S)| \geq w(S)$ . It follows that :  $\mathbf{E}_\xi(\mathcal{M}) = \mathbf{E}_0(\mathcal{M})$ . Since this equality is true is for any ACS  $\mathcal{M}$  then, in particular,  $\mathcal{E}$  is fully exact.

We shall prove that full exactness is a necessary condition of stability of maximal interaction forms. More importantly, the absence of  $r$ -exactness is a symptom of the presence of some cycle of order  $\leq r+2$ . We start by the following :

**Lemma 5.7** *Let  $r \in \mathbb{N}$ ,  $1 \leq r \leq n$ . If  $\mathcal{E}$  is  $r$ -exact and  $E_0(N) = \mathcal{P}_0(A)$ , then for all  $R_N \in L(A)^N : C(E_0, R_N) = \text{Stl}(\mathcal{E}_1, R_N) = \dots = \text{Stl}(\mathcal{E}_r, R_N)$ .*

*Proof.* For any interaction form  $\mathcal{E}$  and any profile  $R_N$  we have  $\text{Stl}(\mathcal{E}_r, R_N) \subset C(E_0, R_N)$ . Assume that  $\mathcal{E}$  is  $r$ -exact and  $E_0(N) = \mathcal{P}_0(A)$ . Let  $a \in A$  be dominated in  $\mathcal{E}_r$  at  $R_N$ . Put  $\varphi(S) = P(a, S, R_N)$  ( $S \in \mathcal{P}_0(N)$ ). Then  $a \notin \rho_{\mathcal{M}}(\varphi)$  and  $\varphi \in \mathcal{E}(\mathcal{M})(\{a\})$  for some  $\mathcal{M} \in \mathfrak{M}_r$  so that  $\varphi \in \mathbf{E}_\xi(\mathcal{M})$ . If for some  $S, T \in \mathcal{M}$ ,  $S \neq T$ ,  $\varphi(S) \cap \varphi(T) \neq \emptyset$ , then  $\varphi(N) = \varphi(S) \cap \varphi(T) \in E_0(N)$  and it follows that  $a$  is dominated in  $E_0$ . If for all  $S, T \in \mathcal{M}$ ,  $S \neq T$ ,  $\varphi(S) \cap \varphi(T) = \emptyset$ , then, by  $r$ -exactness  $\varphi \in \mathbf{E}_0(\mathcal{M}) \cap \mathbf{D}(\mathcal{M})$ . It follows that there exists  $S \in \mathcal{M}$  such that  $\varphi(S) \in E_0(S)$  so that  $a$  is dominated in  $E_0$ .  $\square$

If a maximal interaction form fails to be  $r$ -exact for some  $r \geq 1$ , then it is unstable. More precisely we have the following technical result:

**Lemma 5.8** *Let  $r \in \mathbb{N}$ ,  $1 \leq r \leq n$ . Assume that  $\mathcal{E}$  is maximal. If  $\mathcal{E}$  is not  $r$ -exact then:*

- (i)  $\mathcal{E}_r$  has a cycle of order  $\leq r + 2$ ,
- (ii) If further  $\mathcal{E}$  is superadditive, then there exists some  $R_N \in L(A)^N$  such that  $\text{Stl}(\mathcal{E}_r, R_N) = \emptyset$  and  $|C(E_0, R_N)| = 1$ .

*Proof.* If  $\mathcal{E}$  is not  $r$ -exact, then there exists, an active coalition structure  $\mathcal{M} \in \mathfrak{M}_r$ ,  $a \in A$ ,  $\varphi \in \mathbf{D}(\mathcal{M})$  such that  $\varphi \in \mathcal{E}(\mathcal{M})[a]$ ,  $a \notin \rho_{\mathcal{M}}(\varphi)$ , and for all  $S \in \mathcal{M}$ ,  $\varphi(S) \notin E_0(S)$ . Let  $\mathcal{M} = \{S_1, \dots, S_r\}$ . Let  $T_k = S_k^c$  ( $k = 1, \dots, r$ ),  $B_k = \varphi(S_k)$ ,  $T_0 = N \setminus \cup_1^r T_k$ , and  $B_0 = A \setminus \cup_1^r B_k$ . Since  $E_0$  is maximal we have that  $B_k^c \in E_0(T_k)$  ( $k = 1 \dots, r$ ). Consider the  $r + 2$ -tuple  $((U_0, \varphi_0), \dots, (U_{r+1}, \varphi_{r+1}))$  defined as follows:  $U_0 := B_0 \setminus \{a\}$ ,  $\varphi_0 := \delta_{N, \{a\}}$ ,  $U_k := B_k$ ,  $\varphi_k = \delta_{T_k, B_k^c}$  ( $k = 1, \dots, r$ ),  $U_{r+1} := \{a\}$ ,  $\varphi_{r+1} := \varphi$ . If  $B_0 \setminus \{a\} \neq \emptyset$ , this defines a cycle of order  $r + 2$ . If  $B_0 \setminus \{a\} = \emptyset$  we can remove index 0 and thus have a cycle of order  $r + 1$ .

- (ii) We construct a profile  $R_N = (R_i)_{i \in N}$  with the following properties:

$$\begin{array}{llll} (i \in T_k, k \neq 0) & : A \setminus B_0 \cup B_k & R_i & \{a\} \quad R_i \quad B_0 \setminus \{a\} \quad R_i \quad B_k \\ (i \in T_0) & : \{a\} & R_i & A \setminus B_0 \quad R_i \quad B_0 \setminus \{a\} \end{array}$$

An alternative  $b \in B_k$  where  $k \in \{1, \dots, r\}$  is dominated in  $E_0$  since  $B_k^c \in E_0(T_k)$  and  $B_k^c \subset P(b, T_k, R_N)$ . An alternative  $b \in B_0 \setminus \{a\}$  is dominated in  $E_0$  since by maximality of  $E_0$ ,  $\{a\} \in E_0(N)$  and  $\{a\} \subset P(b, N, R_N)$ . It follows that  $C(E_0, R_N) \subset \{a\}$ . For  $k = 1, \dots, r$ ,  $S_k = \cup_{l \neq k} T_l$ ,  $P(a, S_k, R_N) = \cap_{l \neq k} P(a, T_l, R_N) = \cap_{l \neq k} A \setminus (B_0 \cup B_l) = B_k = \varphi(S_k)$ . So that  $a$  is not a settlement in  $\mathcal{E}_r$ . Since  $\text{Stl}(\mathcal{E}_r, R_N) \subset C(E_0, R_N)$ , it follows that  $\text{Stl}(\mathcal{E}_r, R_N) = \emptyset$ . If furthermore  $E_0$  is superadditive, one can prove that  $C(E_0, R_N) = \{a\}$ .  $\square$

We thus have an easy characterization of  $r$ -exactness when  $E_0$  is maximal and superadditive:

**Proposition 5.9** *Let  $r \in \mathbb{N}$ ,  $1 \leq r \leq n$ . Let  $\mathcal{E}$  be maximal and superadditive. Then  $\mathcal{E}$  is  $r$ -exact if and only if  $\text{Stl}(\mathcal{E}_r, R_N) = C(E_0, R_N)$  for all  $R_N \in L(A)^N$ .*

*Proof.* Since  $E_0$  is maximal, then in particular  $E_0(N) = \mathcal{P}_0(A)$ . If  $\mathcal{E}$  is  $r$ -exact, in view of lemma 5.7,  $\text{Stl}(\mathcal{E}_r, R_N) = C(E_0, R_N)$  for all  $R_N \in L(A)^N$ . The converse follows from lemma 5.8.  $\square$

The following provides necessary and sufficient conditions for stability of  $\mathcal{E}_r$  with maximal  $E_0$ :

**Theorem 5.10** *Let  $r \in \mathbb{N}$ ,  $1 \leq r \leq n$ . Assume that  $E_0$  is maximal.  $\mathcal{E}_r$  is stable if and only if  $E_0$  is stable and  $\mathcal{E}$  is  $r$ -exact. Moreover in this case  $\text{Stl}(\mathcal{E}_r, R_N) = C(E_0, R_N)$  for all  $R_N \in L(A)^N$ .*

Proof. Assume that  $E_0$  is maximal. If  $\mathcal{E}_r$  is stable then in view of lemma 5.8 (ii)  $\mathcal{E}$  is  $r$ -exact. Moreover since  $C(E_0, R_N) \supset \text{Stl}(\mathcal{E}_r, R_N)$  for all  $R_N \in L(A)^N$ ,  $E_0$  is stable. Conversely if  $E_0$  is stable then, in particular  $E_0$  is superadditive and if in addition  $\mathcal{E}$  is  $r$ -exact then  $\text{Stl}(\mathcal{E}_r, R_N) = C(E_0, R_N)$  (Proposition 5.9).  $\square$

Now we deduce necessary conditions for stability of any interaction form  $\mathcal{E}$  with maximal  $E_0$ .

**Theorem 5.11** *Assume that  $E_0$  is maximal. If  $\mathcal{E}$  is stable then:*

- (i)  $E_0$  is stable: that is  $E_0$  is superadditive and subadditive,
- (ii)  $\mathcal{E}$  is fully exact, in particular for all  $1 \leq r \leq n$ , for all  $R_N \in L(A)^N$ , one has :

$$\text{Stl}(\mathcal{E}_r, R_N) = C(E_0, R_N)$$

As a second consequence of lemma 5.8, we obtain the following partial localization of the index in case of unstable  $\mathcal{E}$ :

**Theorem 5.12** *Assume  $\mathcal{E}$  is maximal :*

- (i) If  $\mathcal{E}$  is not regular then  $\sigma(\mathcal{E}) = 2$ ,
- (ii) If  $\mathcal{E}$  is not subadditive or not superadditive then  $\sigma(\mathcal{E}) = 3$
- (iii) If  $\mathcal{E}$  is not  $r$ -exact then  $\sigma(\mathcal{E}) \leq r + 2$ .

*Proof* (i) follows from [5] Theorem 4.9. (ii) is a consequence of Lemma 5.8 (i).

## 5.6 Bipolarity: When the Index is 2

Intuitively a conflictual situation is bipolar when there are two main forces that oppose any settlement. In this short subsection we show the following: whatever is the equilibrium concept that underlies the power system, index 2 is the symptom of the existence of two disjoint coalitions, that can veto any settlement. It is not obvious from the definition of a cycle that a cycle of order 2 marks the existence of two disjoint coalitions with incompatible effectivity power. This will be done under the condition of monotonicity that we now define.

For any  $\mu : \mathcal{P}_0(N) \rightarrow \mathcal{P}_0(N)$  and  $\varphi \in \Phi$  we associate  $\varphi^\mu \in \Phi$  defined by  $\varphi^\mu(S) = \cup\{\varphi(T) : \mu(T) = S\}$ .  $\mu$  is said to be an *inclusion map* if for all  $T \in \mathcal{P}_0(N) : T \subset \mu(T)$ .

**Definition 5.13**  $\mathcal{E}$  is said to be *monotonic* if for any inclusion map  $\mu$  and any  $U \in \mathcal{P}_0(A)$ : if  $\varphi \in \mathcal{E}[U]$  then  $\varphi^\mu \in \mathcal{E}[U]$ .

**Lemma 5.14** *Let  $E$  an effectivity function and let  $\mathcal{E}$  be the interaction form canonically associated to  $E$  (subsection 2.2). Then  $\mathcal{E}$  is monotonic in the sense of the definition 5.13 if and only if  $E$  is monotonic w.r.t. players in the sense of the definition of subsection 2.2.*

*Proof.* Indeed assume that  $\mathcal{E}$  is monotonic. Let  $S \subset T$  and let  $\mu$  be such that  $\mu(S') = S'$  if  $S' \neq S$  and  $\mu(S) = T$ . Then  $\mu$  is an inclusion map. If  $B \in E(S)$  then  $\varphi \equiv \delta_{S,B} \in \mathcal{E}$ . It follows that  $\varphi^\mu = \delta_{T,B} \in \mathcal{E}$  so that  $B \in E(T)$ . Conversely, assume that  $E$  is monotonic w.r.t. players and let  $\mu$  be any inclusion map. If  $\varphi \in \mathcal{E}$ , then for some  $S \in P_0(N)$ ,  $\varphi(S) \in E(S)$  or equivalently  $\varphi' \equiv \delta_{S,\varphi(S)} \in \mathcal{E}$ . Let  $T = \mu(S)$ . Then  $S \subset T$  and  $\varphi'^\mu(S') = \emptyset$  for all  $S' \neq T$ . It follows that  $\varphi'^\mu = \delta_{T,\varphi(S)}$  and  $\varphi(S) \in E(T)$ . We conclude that  $\varphi'^\mu \in \mathcal{E}$ . Now since  $\varphi'^\mu \leq \varphi^\mu$  it follows that  $\varphi^\mu \in E$ .  $\square$

Let  $\mathcal{E}$  is monotonic, let  $\mu$  be an inclusion map,  $\mathcal{M}, \mathcal{M}'$  any ACS such that  $\mu(\mathcal{M}) \subset \mathcal{M}'$ . If  $\varphi \in \mathcal{E}(\mathcal{M})[U]$  then  $\varphi^\mu \in \mathcal{E}(\mathcal{M}')[U]$ .

For any strategic game form  $G$ ,  $\mathcal{E}_\beta^G$  is monotonic.

**Theorem 5.15** *Let  $\mathcal{E}$  be monotonic.*

(i) *The index of  $\mathcal{E}$  is 2 if and only if the index of its derived local effectivity function  $\mathcal{E}_1$  is 2.*

(ii) *If in addition  $\mathcal{E}$  is standard (definition 1.2), then the index of  $\mathcal{E}$  is 2 if and only if the index of its derived (global) effectivity function  $\mathcal{E}_0$  is 2, that is if and only if  $\mathcal{E}$  is not regular.*

*Proof.* Let  $(U_1, \varphi_1), (U_2, \varphi_2)$  be a cycle where wlog, we assume that  $(U_1, U_2)$  is a partition of  $A$ . Let  $\mathcal{M} = [\varphi_1] \cup [\varphi_2]$ . Let  $T_0 = \{i \in N \mid R^i(\varphi_1) \cup R^i(\varphi_2) = \emptyset\}$ . Let  $T_s (s = 1, 2)$  be the set of  $i \in N \setminus T_0$  such that  $U_s \cap (R^i(\varphi_1) \cup R^i(\varphi_2)) = \emptyset$ . By property (ii) of definition 3.1,  $(T_0, T_1, T_2)$  is a partition of  $N$ . For any  $i \in N$ , let  $\mathcal{H}^i := \{S \mid S \ni i\}$ . We have  $i \in T_0$  if and only if  $\mathcal{H}^i \cap \mathcal{M} = \emptyset$ . If  $i \in T_1$  and  $S \in \mathcal{H}^i$  then  $\varphi_1(S) \cup \varphi_2(S) \subset U_2$ . Similarly if  $j \in T_2$  and  $S \in \mathcal{H}^j$  then  $\varphi_1(S) \cup \varphi_2(S) \subset U_1$ . It follows that  $\mathcal{H}^i \cap \mathcal{H}^j$  contains only coalitions  $S$  such that  $\varphi_1(S) \cup \varphi_2(S) = \emptyset$  or put equivalently  $\mathcal{H}^i \cap \mathcal{H}^j \cap \mathcal{M} = \emptyset$  or put otherwise: for any  $S \in \mathcal{M}$  either  $S \subset T_1$  or  $S \subset T_2$ . Moreover by property (ii) of definition 3.1, if  $S \subset T_1$  then  $\varphi_2(S) \cap (U_1 \cup U_2) = \emptyset$ . Since  $U_1 \cup U_2 = A$ , it follows that  $\varphi_2(S) = \emptyset$ . Similarly if  $S \subset T_2$  then  $\varphi_1(S) = \emptyset$ . Let  $\mu : \mathcal{P}_0(N) \rightarrow \mathcal{P}_0(N)$  defined by  $\mu(S) = T_k$  if  $S \subset T_k$  ( $k = 1, 2$ ) and  $\mu(S) = N$  if  $S \notin \mathcal{M}$ . Since  $\mathcal{E}$  is monotonic:  $\varphi_k^\mu \in \mathcal{E}[U_k]$  ( $k = 1, 2$ ). Moreover if  $S \neq T_k$ ,  $\varphi_k^\mu(S) = \emptyset$ . It follows that  $\varphi_k^\mu(T_k) \in E_1[U_k](T_k)$  ( $(U_1, \varphi_1^\mu), (U_2, \varphi_2^\mu)$ ) is a 2-cycle for  $\mathcal{E}_1$ . Since  $\varphi_1^\mu(T_1) \subset U_2$  and  $\varphi_2^\mu(T_2) \subset U_1$ , we have  $U_2 \in E_1[U_1](T_1)$  and  $U_1 \in E_1[U_2](T_2)$ . If we assume in addition that  $\mathcal{E}$  is locally effective, and satisfies the sheaf property, then  $U_1 \in E_0(S_2)$  and  $U_2 \in E_0(S_1)$ , so that  $\mathcal{E}_0$  is not regular.  $\square$

## 6 Stability index of strategic game forms

In this section we adopt notations of subsections 2.1 and 1.2. We consider a *game form*  $G = (X_1, \dots, X_n, A, g)$ . The classical effectivity function  $E_\beta^G$ , is defined by :

$$E_\beta^G(S) = \{B \in \mathcal{P}_0(A) \mid \forall y_N \in X_N, \exists x_S \in X_S : g(x_S, y_S) \in B\} \quad (31)$$

Clearly  $E_\beta^G$  is an effectivity function (see subsection 2.2). Actually  $E_\beta^G$  is the (global) effectivity function derived from  $\mathcal{E}_\beta^G$  (see subsections 2.1, and equation 21). The  $\alpha$ -effectivity function is defined by:

$$E_\alpha^G(S) = \{B \in \mathcal{P}_0(A) \mid B^c \notin \mathcal{E}_\beta^G(S)\} \quad (32)$$

It is easy to see that  $E_\beta^G$  is maximal, that  $E_\alpha^G$  is superadditive (hence regular). Moreover  $\mathcal{E}_\beta^G$  is monotonic (definition 5.13).  $G$  is said to be *tight* if  $E_\alpha^G = \mathcal{E}_\beta^G$ .  $G$  is tight if and only if  $E_\alpha^G$  is maximal or equivalently if  $E_\beta^G$  regular.

In what follows for  $r \geq 1$ ,  $(\mathcal{E}_\beta^G)_r$  is denoted simply  $\mathcal{E}_r^G$ , similarly if  $\mathfrak{F}$  is a federation  $(\mathcal{E}_\beta^G)_{\mathfrak{F}}$  will be denoted  $\mathcal{E}_{\mathfrak{F}}^G$ . For the sake of consistency in our notation and vocabulary, we shall denote by  $C_{\mathfrak{F}}(G, R_N)$ , and we shall call the  $\mathfrak{F}$ -core of  $G$  at  $R_N$ , the set  $\cap_{M \in \mathfrak{F}} EO(G, R_N)$ . In summary, while equilibria have to do with some ACS and the projection of  $\mathcal{E}_\beta^G$  on that ACS, the cores have to do with some federation and the restriction of  $\mathcal{E}_\beta^G$  on that federation. Note that  $E_0$  in the preceding section corresponds to  $E_\beta^G$  in the present section.

**Definition 6.1** Let  $\mathcal{M}$  be an ACS. The  $\mathcal{M}$ -*stability index* of  $G$  is the stability index of  $\mathcal{E}_\beta^G(\mathcal{M})$ . It will be denoted  $\sigma(G, \mathcal{M})$ . Similarly if  $\mathfrak{F}$  is a federation, the  $\mathfrak{F}$ -*core stability index* of  $G$  is the stability index of  $\mathcal{E}_{\mathfrak{F}}^G$ . It will be denoted by  $\sigma(G, \mathfrak{F})$ .

Stability index of  $\mathcal{E}_0^G$  (that is the stability index of  $G$  relative to the  $\beta$ -core) and the stability index of  $\mathcal{E}_1^G$  (that is the stability index of  $G$  relative to the local core) have been the object of [5] (Theorem 4.9 and Theorem 4.15). We start by recalling the result for the  $\beta$ -core:

**Theorem 6.2** *Let  $G$  be a game form. Let  $\sigma_0$  be its index for the  $\beta$ -core solution. Then:*

- (i)  $\sigma_0 = 2$  if and only if  $G$  is not tight,
- (ii)  $\sigma_0 = 3$  if  $G$  is tight but  $E_\beta^G$  is not subadditive
- (iii)  $\sigma_0 = +\infty$  if  $G$  is tight and  $E_\beta^G$  subadditive.

In the rest of this section we shall establish some facts about the stability index for the major equilibrium concepts namely Nash and strong Nash equilibrium. For strong solvability we have the following:

**Theorem 6.3** (i) *If  $G$  is strongly solvable then  $G$  is tight,  $E_\beta^G$  is superadditive and subadditive and  $\mathcal{E}_\beta^G$  is fully exact.*

(ii) *If  $n = 2$  then  $G$  is strongly solvable if and only if  $E_\beta^G$  is regular and  $\mathcal{E}_\beta^G$  is 2-exact.*

*Proof.*  $\mathcal{E}_\beta^G$  is maximal so that this result is a straightforward consequence of theorem 5.10.  $\square$

**Remark 6.4** In [4] example 2.3, a 2-player game form is given such that  $\text{Stl}(\mathcal{E}_1^G, R_N) \neq \text{Stl}(\mathcal{E}_0^G, R_N) \equiv C(E_\beta^G, R_N)$ . In example 2.4, a 2-player game form is given such that  $\text{Stl}(\mathcal{E}_2^G, R_N) \neq \text{Stl}(\mathcal{E}_1^G, R_N)$ , and in example 2.5 a 3-player game form is given where  $\text{Stl}(\mathcal{E}_2^G, R_N)$  is strictly larger than the corresponding strong outcome set denoted  $SEO(G, R_N)$ . It would be interesting to exhibit an example of an  $n$ -player game form (where necessarily  $n \geq 3$ ) such that  $G$  is strongly solvable and  $SEO(G, R_N) \equiv \text{Stl}(\mathcal{E}_\beta^G, R_N) \neq \text{Stl}(\mathcal{E}_n^G, R_N)$  for some  $R_N$ . By Theorem 6.3 this is equivalent to find  $G$  such that  $G$  is strongly solvable and  $SEO(G, R_N) \neq C(E_\beta^G, R_N)$  for some  $R_N$ .

In case of strong Nash instability, we can determine upper bounds for the strong Nash index.

**Theorem 6.5** *Let  $G$  be a game form. Let  $\sigma$  be its index for strong Nash Equilibrium then:*

- (i)  $\sigma = 2$  if and only if  $G$  is not tight,
- (ii)  $\sigma = 3$  if  $G$  is tight but  $E_\beta^G$  is not subadditive,
- (iii)  $\sigma \leq r + 2$  if  $\mathcal{E}_\beta^G$  is not  $r$ -exact ( $1 \leq r \leq n$ ).

*Proof.* The result is a straightforward consequence of theorem 5.12 and the remark that the interaction form  $\mathcal{E}_\beta^G$  is monotonic and in case of tightness it is superadditive.  $\square$

In case of bipolarity one has the following corollary of theorem 5.15:

**Corollary 6.6** *Let  $G$  be a strategic game form. If the index is 2 for some  $\mathcal{M}$ -equilibrium concept, where  $\emptyset \neq \mathcal{M} \subset \mathcal{P}_0(N)$ , then the  $\beta$ -effectivity function of  $G$  has index 2, that is the game form is not tight.*

*Proof.*  $\mathcal{E}_\beta^G$  is a standard monotonic interaction form. Therefore the result follows from Theorem 5.15.  $\square$

Now we give a localization of the index for some classes of games, that can be obtained as corollaries from known results in the literature.

**Theorem 6.7** *The Nash stability index of a two-player game form is either 2 or  $+\infty$*

This is a corollary of the fact that for these game forms Nash solvability is equivalent to tightness. (See Gurvich [9, 11] or Abdou [2]). If we consider the class of rectangular game forms i.e. such that for any  $a \in A$ ,  $g^{-1}(a) = \prod_{i=1}^n Y_i$ , for some  $Y_i \subset X_i$ , ( $i = 1, \dots, n$ ) (Gurvich [10] and Abdou [2, 3]) one has a similar characterization for Nash solvability by tightness. However *it does not follow* that the Nash stability index of a rectangular game form is 2 in case of Nash instability. The only property that can be asserted in this case is that the  $\beta$ -core index of such a game form is 2. The strong Nash index for rectangular game forms has a simple characterization:

**Theorem 6.8** *The strong Nash index of any rectangular game form is either 2 or 3 or  $+\infty$ .*

Proof: In view of theorem 4.7 of Abdou [4] any rectangular game form  $G$  such that  $\mathcal{E}_\beta^G$  is 1-exact is essentially a one-player game form. It follows that a rectangular game form is strongly solvable if and only if it is 1-exact and in this case the strong Nash index is  $+\infty$ . If  $G$  is not strongly solvable,  $\mathcal{E}_\beta^G$  is not 1-exact, then by theorem 6.5 (iii) the index is less than  $1 + 2 = 3$ , and in fact is equal 3 if  $G$  is tight (that is Nash solvable).  $\square$

We end this section by presenting some details about the interaction forms for two examples presented earlier.

**Example 6.9** Take the Majority game form of example 2.3. One can observe that  $G$  is tight and 3-exact. However it is not subadditive. Therefore the effectivity function  $E_\beta^G$  is not stable so that the  $\beta$ -core index of  $G$  is 3. The universal interaction form of this game, namely  $\mathcal{E}_\beta^G \equiv \mathcal{E}_\beta^G(\mathcal{P}_0(N))$  has index 3 too: if  $\mathcal{E}_\beta^G$  had some cycle of order 2, it would give rise to some cycle of order 2 for  $E_\beta^G$ . Therefore the strong Nash index of  $G$  is 3.

**Example 6.10** Take the Unanimity game form of example 2.4. Then  $\mathcal{E}_\alpha^G[a]$  and  $\mathcal{E}_\beta[a]$  have been determined for any  $a \in X$ . Here we shall determine  $\mathcal{E}_\alpha^G(\mathcal{M})[0]$  and  $\mathcal{E}_\beta(\mathcal{M})[0]$  where  $\mathcal{M}$  is in the class  $\mathfrak{M}$  of admissible ACS. Let  $\mathcal{M} = \{S_k : k = 1, \dots, r\}$  ( $r \geq 1$ ) be an admissible ACS. We associate to  $\mathcal{M}$  the sets:

$$\Delta_n := \{(x_1, \dots, x_n) \in X^n \mid x_1 = \dots = x_n\} \quad (33)$$

$$\mathcal{M}_1 = \{S \in \mathcal{M} \mid |S^c| = 1\} \quad (34)$$

$$J(\mathcal{M}) = \{i \mid N \setminus \{i\} \in \mathcal{M}_1\} \quad (35)$$

$$\Gamma_0(\mathcal{M}) := \{\varphi \mid 0 \in \rho_{\mathcal{M}}(\varphi)\} \quad (36)$$

$$\Psi(\mathcal{M}) := \{\varphi \mid \exists S \in \mathcal{M} : \varphi(S) = X\} \quad (37)$$

$$\Lambda(\mathcal{M}) = \{\varphi \mid \exists a \in X, \forall S \in \mathcal{M} : \varphi(S) = X \setminus \{a\}\} \quad (38)$$



We first treat the case where  $\mathcal{M}^* = \{S_1^c, \dots, S_r^c\}$  is a partition of  $N$ , that is  $\cap_{S \in \mathcal{M}} S = \emptyset$ . In particular  $r \geq 2$ .

**Claim:**  $\mathcal{E}_\alpha(\mathcal{M})[0] \equiv \mathcal{E}_\alpha^G(\mathcal{M})[0]$  and  $E_\beta^G(\mathcal{M})[0] \equiv E_\beta(\mathcal{M})[0]$  are given by the following:

If  $\mathcal{M}_1 = \mathcal{M}$  :

$$\mathcal{E}_\alpha(\mathcal{M})[0] = \{\psi \in \Phi \mid \exists x \in X^n \setminus \Delta, \forall i \in N : \{0, x_i\} \subset \psi(N \setminus \{i\})\} \quad (39)$$

$$\mathcal{E}_\beta(\mathcal{M})[0] = \Gamma_0(\mathcal{M}) \cup \Psi(\mathcal{M}_1) \cup \Lambda(\mathcal{M}) \quad (40)$$

If  $\mathcal{M}_1 \neq \mathcal{M}$  :

$$\mathcal{E}_\alpha(\mathcal{M})[0] = \{\psi \mid \forall S \in \mathcal{M} : 0 \in \psi(S) \text{ and } \forall S \in \mathcal{M}_1 : |\psi(S)| \geq 2\} \quad (41)$$

$$\mathcal{E}_\beta(\mathcal{M})[0] = \Gamma_0(\mathcal{M}) \cup \Psi(\mathcal{M}_1) \quad (42)$$

*Proof of the claim.* The formula for  $\mathcal{E}_\alpha(\mathcal{M})[0]$  is straightforward. Let  $\varphi \in \Phi$  such that  $\rho_\mathcal{M}(\varphi) \subset X$ . Then if  $\mathcal{M}_1 = \mathcal{M}$ , we have  $J(\mathcal{M}) = N$  and :

$$\varphi \notin \mathcal{E}_\beta(\mathcal{M})[0] \Leftrightarrow \varphi^* \in \mathcal{E}_\alpha(\mathcal{M})[0] \quad (43)$$

$$\Leftrightarrow \exists x \in X^n \setminus \Delta_n, \forall i \in J(\mathcal{M}), x_i \in \varphi^*(N \setminus \{i\}) \quad (44)$$

$$\Leftrightarrow \prod_{i \in N} (X \setminus \varphi(N \setminus \{i\})) \cap (X^n \setminus \Delta_n) \neq \emptyset \quad (45)$$

It follows that when  $\varphi$  is not in  $\Gamma_0(\mathcal{M})$ ,  $\varphi$  is in  $\mathcal{E}_\beta(\mathcal{M})[0]$  if and only if  $\prod_{i \in N} (X \setminus \varphi(N \setminus \{i\})) \subset \Delta_n$ . Then either  $\prod_{i \in N} (X \setminus \varphi(N \setminus \{i\}))$  is empty and  $\varphi \in \Psi(\mathcal{M})$  or there exists  $a \in X$ , such that  $X \setminus \varphi(N \setminus \{i\}) = \{a\}$  for all  $i \in N$  and  $\varphi$  is in  $\Lambda(\mathcal{M})$ .

When  $\mathcal{M}_1 \neq \mathcal{M}$ , we have  $J(\mathcal{M}) \neq N$ , so that the RHS of formula (45) is replaced by  $\prod_{i \in J(\mathcal{M})} (X \setminus \varphi(N \setminus \{i\})) \cap X^{J(\mathcal{M})} \neq \emptyset$ . It follows that when  $\varphi$  is not in  $\Gamma_0(\mathcal{M})$ ,  $\varphi$  is in  $\mathcal{E}_\beta(\mathcal{M})[0]$  if and only if  $\prod_{i \in J(\mathcal{M})} (X \setminus \varphi(N \setminus \{i\})) = \emptyset$ , so that we conclude that  $\varphi$  is in  $\Psi(\mathcal{M}_1)$ . Remark that when  $\mathcal{M}_1$  is empty, so is  $\Psi(\mathcal{M}_1)$ .  $\square$

If  $\mathcal{M}^*$  is not a partition of  $N$ , then let  $S_0 = \cup_k^r S_k^c$ . Put  $\mathcal{M}' = \mathcal{M} \cup \{S_0\}$ .  $\mathcal{E}_\beta(\mathcal{M})[0]$  is then calculated as the projection of  $\mathcal{E}_\beta(\mathcal{M}')[0]$  on  $\mathcal{M}$ . One obtains the general formula whether  $\mathcal{M}^*$  is a partition or not:

**Proposition 6.11** *For any  $\mathcal{M} \in \mathfrak{M}$ ,  $\mathcal{M} \neq \{N\}$  one has :*

$$\begin{aligned} \mathcal{E}_\beta(\mathcal{M})[0] &= \Gamma_0 \cup \Psi(\mathcal{M}_1), & \text{if } 0 \leq |\mathcal{M}_1| < n. \\ &= \Gamma_0(\mathcal{M}) \cup \Psi(\mathcal{M}_1) \cup \Lambda(\mathcal{M}) & \text{if } |\mathcal{M}_1| = n \end{aligned}$$

In particular taking  $\mathcal{M} = \{S\}$ , where  $S \in \mathcal{P}_0(N)$   $S \neq N$ , we can deduce:

$$\begin{aligned} E_\beta(S) &= \{B \in \mathcal{P}_0(A) \mid 0 \in B\} && \text{if } 1 \leq |S| \leq n-2 \\ &= \{B \in \mathcal{P}_0(A) \mid 0 \in B\} \cup \{X\} && \text{if } |S| = n-1 \\ &= \mathcal{P}_0(A) && \text{if } S = N \end{aligned}$$

In order to study exactness we need  $\mathbf{E}_0(\mathcal{M}) \equiv \mathcal{E}_\beta(\mathcal{M})[A] \cap \{\varphi \mid \rho_{\mathcal{M}}(\varphi) \neq A\}$  and  $\mathbf{E}_\xi(\mathcal{M})$ . They are given by the following:

$$\mathbf{E}_0(\mathcal{M}) = \Gamma_0(\mathcal{M}) \cup \Psi(\mathcal{M}_1) \cap \{\varphi \mid \rho_{\mathcal{M}}(\varphi) \neq A\} \quad (46)$$

$$\mathbf{E}_\xi(\mathcal{M}) = \mathcal{E}_\beta(\mathcal{M})[0] \cap \{\varphi \mid \rho_{\mathcal{M}}(\varphi) \neq A\} \quad (47)$$

By writing the intersection of these sets with  $\mathbf{D}(\mathcal{M})$ , one sees that  $\mathcal{E}_\beta$  is  $\mathcal{M}$ -exact for any  $\mathcal{M} \in \mathfrak{M}$  or equivalently  $\mathcal{E}_\beta$  is fully exact. However note that since  $X \in E_\beta(S)$  and  $\{0\} \in E_\beta(S^c)$  if  $|S| = n-1$ , it follows that  $E_\beta$  is not regular; we conclude that the  $\beta$ -core index and consequently the strong Nash index of a the unanimity game is 2.

## 7 Concluding remarks

The model of interaction form as a description of power distribution of a set of agents  $N$  over a set of alternatives  $A$  encompasses aspects of both cooperative and strategic models. The settlement set defined at a preference profile reflects the outcomes that may emerge given the power system. Any game form in the context of a classical solution (equilibrium concept or core) gives rise to some interaction form. Interactive forms defined on the same sets of agents and alternatives can be compared with each other. In particular they can be compared with respect to their stability, a main issue in political science and social choice. Stability is proven to be equivalent to acyclicity. Solvability of strategic game forms is thus reduced to absence of cycles. In order to describe the properties of unstable interaction forms, we introduced a graded notion of exactness. A failure of exactness of some order for a maximal interaction form is a sign of the existence of some cycle. We can thus localize the stability index in many cases. Many questions about the stability index are still open. Moreover a finer analysis of the structure of instability requires the definition of a typology of cycles that goes beyond the notion of index.

## References

- [1] Abdou, J., 1982. Stabilité de la fonction veto, Cas du veto maximal. Math. Sci. Hum. 80, 39-63.
- [2] Abdou, J., 1995. Nash and strongly consistent two-player game forms. Int. J. Game Theory 24, 345-356.

- [3] Abdou, J., 1998. Rectangularity and tightness: a normal form characterization of perfect information extensive game forms. *Math. of oper. research* 23, No 3, 553-567.
- [4] Abdou, J., 2000. Exact stability and its applications to strong solvability. *Math. Social Sciences* 39, 263 - 275.
- [5] Abdou, J., 2009. A stability index for local effectivity functions.
- [6] Abdou, J. and Keiding H., 1991. Effectivity functions in social choice. Kluwer Academic Publishers, Dordrecht.
- [7] Abdou, J. and Keiding H., 2003. On necessary and sufficient conditions for solvability of game forms. *Math. Social Sciences* 46, 243 - 260.
- [8] Greenberg, J. 1990. *The Theory of Social Situations: An Alternative Game-Theoretic Approach*, Cambridge University Press.
- [9] Gurvich, V.A., 1975. Solvability in pure strategies. *Zh. Vychisl. Mat. i Mat. Fiz.* 15, 358-371; English transl. in *USSR Comput. Math. and Math. Phys.* 15.
- [10] Gurvich, V.A., 1978. Application of Boolean functions and contact schemes in game theory. Candidate's dissertation Moskow, Fiz-Tekhn. Inst. Russian.
- [11] Gurvich, V.A., 1989. Equilibrium in pure strategies. *Soviet Math. Dokl.* 38, 597-602.
- [12] Keiding, H., 1985. Necessary and sufficient conditions for stability of effectivity functions, *Int. J. Game Theory* 14, 93 - 101.
- [13] Moulin, H. and Peleg, B., 1982. Cores of effectivity functions and implementation theory. *J. Math. Econ.* 10, 115-162.
- Nakamura, K. (1979), The vetoers in a simple game with ordinal preferences, *Int. J. Game Theory* 8, issue 1, 55-61
- [14] Peleg, B., 1984. *Game theoretic analysis of voting in committees*. Cambridge University Press, Cambridge.
- [15] Peleg, B., 1998. Effectivity functions, game forms, games, and rights. *Social Choice and Welfare* 15: 67-80.
- [16] Peleg, B., 2004. Representation of Effectivity functions by acceptable game forms: a complete characterization. *Mathematical Social Sciences* 47, 275-287.

- [17] Peleg, B. and Peters, H., 2008. Nash consistent representations of effectivity functions through lottery models. Forthcoming in *Games and Economic Behavior* .
- [18] Peleg, B. and Winter, E., 2002. Constitutional implementation. *Review of Economic Design* 7, 187–204.
- [19] Rosenthal, R.W., 1972. Cooperative games in effectiveness form, *J. Econ. Theory* 5 88101.
- [20] Vannucci, S., 2002. Effectivity Functions and Stable Governance Structures. *Annals of Operations Research* 109, 99127.
- [21] Vannucci, S., 2008. A coalitional game-theoretic model of stable government forms with umpires. *Rev. Econ. Design* 12:3344.